Option Pricing, Portfolio Optimization, and Stochastic Calculus

Holger Kraft

Goethe University

Faculty of Economics and Business Administration

First Version: July 29, 2005

This Version: December 7, 2015

Preface

These are notes that I have used to *supplement* my teaching of continuous-time option pricing and portfolio theory. The focus is on a rigorous treatment of a Brownian setting. The notes are supposed to help students who want or need to learn more about the formal concepts behind option pricing theory. This material is challenging. Therefore, the first two sections, Section 1 and 2, review the pricing of derivatives in a simple discrete-time model with finite state space. These sections should be accessible without particular knowledge in probability theory. Nevertheless, the main insights (e.g. replication, deflator, equivalent martingale measure) carry over to the continuous-time framework that is introduced in Section 3 and at the beginning of Section 4. Section 4 studies option pricing in continuous time. Here some background in probability theory and stochastic calculus is needed. The necessary results are provided in Section 3. These sections should be accessible if you are familiar with the material of textbooks like "Probability Essentials" by Jean Jacod and Philip Protter. Section 5 gives a brief introduction to continuous-time portfolio theory.

The notes are not intended to be an economic textbook on option pricing. They are written in a "definition, proposition, proof" style. I have tried to give some economic intuition, but there could be much more (sorry!) and in my lectures there is much more. To learn more about the economics of option pricing or asset pricing in general, you might consider books like Merton (1990), Duffie (2001), Cochrane (2001), and Munk (2013) and the references therein.

I have tried to make the notes self-contained, but sometimes I need to refer to other sources that can be found at the end of the notes. For omitted technical results, my main source of reference is the book by Korn/Korn (2001). For instance, a proof of the martingale representation theorem can be found there. Of course, there are a lot of other good books and the list of references does not attempt to be complete.

Contents

1 Introduction

1.1 What is an Option?

- There exist two main variants: call and put.
- A European call gives the owner the right to buy some asset (e.g. stock, bond, electricity, corn ...) at a future date (maturity) for a fixed price (strike price).
- Payoff: $\max\{S(T) - K; 0\}$, where $S(T)$ is the asset price at maturity T and K is the strike price.
- An American call gives the owner the right to buy some asset during the whole life-time of the call.
- Put: replace "buy" with "sell"
- Since the owner has the right but not the obligation, this right has a positive value.

1.2 Fundamental Relations

Consider an option written on an underlying which pays no dividends etc. during the life-time of the option. In the following we assume that the underlying is a stock. The time-t call and put prices are denoted by $Call(t)$ and $Put(t)$. The time-t price of a zero-coupon bond with maturity T is denoted by $p(t, T)$.

Proposition 1.1 (Bounds)
(i) For a European call we get: $S(t) \geq Call(t) \geq \max\{S(t) - Kp(t,T); 0\}$

(ii) For a European put we have: $Kp(t,T) \geq Put(t) \geq \max\{Kp(t,T) - S(t); 0\}$

Proof. (i) The relation $S(t) \geq Call(t)$ is obvious because the stock can be interpreted as a call with strike 0. We now assume that $Call(t) < \max\{S(t) - Kp(t,T); 0\}$. W.l.o.g. $S(t) - Kp(t,T) \geq 0$. Consider the strategy
- buy one call,
- sell one stock,
- buy K zeros with maturity T,

leading to an initial inflow of $S(t) - Call(t) - Kp(t,T) > 0$. At time T, we distinguish two scenarios:

1. $S(T) > K$: Then the value of the strategy is $-S(T) + S(T) - K + K = 0$.
2. $S(T) \leq K$: Then the value of the strategy is $-S(T) + 0 + K \geq 0$.

Hence, the strategy would be an arbitrage opportunity.

(ii) The relation $Kp(t,T) \geq Put(t)$ is valid because otherwise the strategy
- sell one put,
- buy K zeros with maturity T

would be an arbitrage strategy. The second relation can be proved similarly to the relation for the call in (i). □

Proposition 1.2 (Put-Call Parity)
The following relation holds: $Put(t) = Call(t) - S(t) + Kp(t,T)$

Proof. Consider the strategy
- buy one put,
- sell one call,
- buy one stock,
- sell K zeros with maturity T.

It is easy to check that the time-T value of this strategy is zero. Since no additional payments need to be made, no arbitrage dictates that the time-t value is zero as well implying the desired result. □

Remark. This is a very important relationship because it ties together European call and put prices.

Proposition 1.3 (American Call)
Assume that interest rates are positive. Early exercise of an American call is not optimal, i.e. $Call^{am}(t) = Call(t)$.

Proof. The following relations hold

$$Call^{am}(t) \geq Call(t) \overset{(*)}{\geq} \max\{S(t) - Kp(t,T); 0\} \geq S(t) - K$$

where $(*)$ holds due to Proposition 1.1 (i). Hence, the American call is worth more alive than dead. □

Remark.
- Unfortunately, this result does not hold for American puts.
- If dividends are paid, then the result for the American call breaks down as well.

Proposition 1.4 (American Put)
Assume that interest rates are positive. Then the following relations hold:

(i) $Call^{am}(t) - S(t) + Kp(t,T) \leq Put^{am}(t) \leq Call^{am}(t) - S(t) + K$

(ii) $Put(t) \leq Put^{am}(t) \leq Put(t) + K(1 - p(t,T))$.

Proof. (i) The first inequality follows from the put-call parity and Proposition 1.3. For the second one assume that $Put^{am}(t) > Call^{am}(t) - S(t) + K$ and consider the following strategy:
- sell one put,
- buy one call,
- sell one stock,
- invest K Euros in the money market account.[1]

[1] The time-t value of the money market account is denoted by M(t). By convention, $M(0) = 1$.

By assumption, the initial value of this strategy is strictly positive. For $t \in [0, T]$ we distinguish to scenarios:

1. $S(t) > K$: Then the value of the strategy is

$$0 + (S(t) - K) - S(t) + KM(t) = K(M(t) - 1) \geq 0$$

since interest rates are positive and thus $M(t) \geq 1$.

2. $S(t) \leq K$: Then the value of the strategy is

$$-(K - S(t)) + 0 - S(t) + KM(t) = K(M(t) - 1) \geq 0.$$

Hence, the strategy would be an arbitrage opportunity.

(ii) We have

$$\begin{aligned} Put(t) &\leq Put^{am}(t) \leq Call^{am}(t) - S(t) + K = Call(t) - S(t) + K \\ &= Put(t) + S(t) - Kp(t, T) - S(t) + K = Put(t) + K(1 - p(t, T)). \end{aligned}$$

□

Remark. If interest rates are zero, then $Put(t) = Put^{am}(t)$.

2 Discrete-time State Pricing with Finite State Space

2.1 Single-period Model

2.1.1 Description of the Model

- We consider a single-period model starting today ($t = 0$) and ending at time $t = T$. Without loss of generality $T = 1$.
- Uncertainty at time $t = 1$ is represented by a finite set $\Omega = \{\omega_1, \ldots, \omega_I\}$ of states, one of which will be revealed as true. Ω is said to be the state space.
- Each state $\omega_i \in \Omega$ can occur with probability $P(\omega_i) > 0$. This defines a probability measure $P : \Omega \to (0, 1)$.
- At time $t = 0$ one can trade in $J + 1$ assets. The time-0 prices of these assets are described by the vector $S' = (S_0, S_1, \ldots, S_J)$. The state-dependent payoffs of the assets at time $t = 1$ are described by the payoff matrix

$$X = \begin{pmatrix} X_0(\omega_1) & \cdots & X_J(\omega_1) \\ X_0(\omega_2) & \cdots & X_J(\omega_2) \\ \vdots & \ddots & \vdots \\ X_0(\omega_I) & \cdots & X_J(\omega_I) \end{pmatrix}.$$

- To shorten notations, $X_{ij} := X_j(\omega_i)$.
- The first asset is assumed to be riskless, i.e. $X_{i0} = X_{k0} = const.$ for all i, $k \in \{1, \ldots, I\}$.
- Therefore, at time $t = 0$ there exists a riskless interest rate given by

$$r = \frac{X_0}{S_0} - 1.$$

- An investor buys and sells assets at time $t = 0$. This is modeled via a trading strategy $\varphi = (\varphi_0, \ldots, \varphi_J)'$, where φ_j denotes the number of asset j which the investor holds in his portfolio (combination of assets).

Definition 2.1 (Arbitrage)
We distinguish two kinds of arbitrage opportunities:

(i) A free lunch is a trading strategy φ with

$$S'\varphi < 0 \quad \textit{and} \quad X\varphi \geq 0.$$

(ii) A free lottery is a trading strategy φ with

$$S'\varphi = 0 \quad \textit{and} \quad X\varphi \geq 0 \quad \textit{and} \quad X\varphi \neq 0.$$

Assumptions:

(a) All investors agree upon which states cannot occur at time $t = 1$.

(b) All investors estimate the same payoff matrix (homogeneous beliefs).

(c) At time $t = 0$ there is only one market price for each security and each investor can buy and sell as many securities as desired without changing the market price (price taker).

(d) Frictionless markets, i.e.
- assets are divisible,
- no restrictions on short sales,[2]
- no transaction costs,
- no taxes.

(e) The asset market contains no arbitrage opportunities, i.e. there are neither free lunches nor free lotteries.

2.1.2 Arrow-Debreu Securities

To characterize arbitrage-free markets, the notion of an Arrow-Debreu security is crucial:

Definition 2.2 (Arrow-Debreu Security)
An asset paying one euro in exactly one state and zero else is said to be an Arrow-Debreu security (ADS). Its price is an Arrow-Debreu price (ADP).

Remarks.
- In our model, we have as many ADS as states at time $t = 1$, i.e. we have I Arrow-Debreu securities.
- The i-th ADS has the payoff $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)'$, where the i-th entry is 1. The corresponding Arrow-Debreu price is denoted by π_i.
- Unfortunately, in practice ADSs are not traded.

We can now prove the following theorem characterizing markets which do not contain arbitrage opportunities.

Theorem 2.1 (Characterization of No Arbitrage)
Under assumptions (a)-(d) the following statements are equivalent:

(i) The market does not contain arbitrage opportunities (Assumption (e)).

[2] A short sale of an assets is equivalent to selling an asset one does not own. An investor can do this by borrowing the asset from a third party and selling the borrowed security. The net position is a cash inflow equal to the price of the security and a liability (borrowing) of the security to the third party. In abstract mathematical terms, a short sale corresponds to buying a negative amount of a security.

(ii) There exist strictly positive Arrow-Debreu prices, i.e. $\pi_i > 0$, $i = 1, \ldots, I$, *such that*

$$S_j = \sum_{i=1}^{I} \pi_i \cdot X_{ij}.$$

Proof. Suppose (ii) to hold. We consider a strategy φ with $X\varphi \geq 0$ and $X\varphi \neq 0$. Then

$$S'\varphi \overset{(ii)}{=} \pi' X\varphi \overset{(*)}{>} 0,$$

where $(*)$ holds because $\pi > 0$, $X\varphi \geq 0$, and $X\varphi \neq 0$. For this reason, there are no free lotteries. The argument to exclude free lunches is similar. The implication "(i)⇒(ii)" follows from the so-called Farkas-Stiemke lemma (See Mangasarian (1969, p. 32) or Dax (1997)). □

Remarks.

- The ADPs are the solution of the following system of linear equations:

$$S = X'\pi.$$

- The hard part of the proof is actually the proof of implication "(i)⇒(ii)". For brevity, this proof is omitted here.

Given our market with prices S and payoff matrix X, it is now of interest how to price other assets.

Definition 2.3 (Contingent Claim)
In the single-period model, a contingent claim (derivative, option) is some additional asset with payoff $C \in I\!R^I$.

Definition 2.4 (Replication)
Consider a market with payoff matrix X*. A contingent claim with payoff* C *can be replicated (hedged) in this market if there exists a trading strategy* φ *such that*

$$C = X\varphi.$$

Remark. If an asset with payoff C is replicable via a trading strategy φ, by the no arbitrage assumption its price S_C has to be ("law of one price")

$$S_C = \sum_{j=1}^{J} \varphi_j \cdot S_j.$$

Definition 2.5 (Completeness)
An arbitrage-free market is said to be complete if every contingent claim is replicable.

Theorem 2.2 (Characterization of a Complete Market)
Assume that the financial market contains no arbitrage opportunities. Then the market is complete if and only if the Arrow-Debreu prices are uniquely determined, i.e.

$$Rank(X) = I.$$

Proof.

$$\begin{aligned} & \pi \text{ is unique.} \\ \iff\ & X'\pi = S \text{ has a unique solution.} \\ \iff\ & \text{Rank } (X) = I \\ \iff\ & \text{Range } (X) = I\!R^I \\ \iff\ & \text{Every contingent claim is replicable.} \end{aligned}$$

□

Remarks.

- The market is complete if the number of assets with linearly independent payoffs is equal to the number of states.
- If the number of assets is smaller than the number of states ($J < I$), the market cannot be complete.

2.1.3 Deflator

Main result from the previous subsection: In an arbitrage-free market, there exist strictly positive ADPs such that

$$S_j = \sum_{i=1}^{I} \pi_i \cdot X_{ij}.$$

In this subsection: Using the physical probabilities $P(\omega_i)$, $i = 1, \ldots, I$, arbitrage-free asset prices can be represented as discounted expected values of the corresponding payoffs.

Definition 2.6 (Deflator)
In an arbitrage-free market, a deflator H is defined by

$$H = \frac{\pi}{P}.$$

Remarks.

- Since the market is arbitrage-free, there exists at least one vector of ADSs π, but π and thus the deflator H are only unique if the market is complete.
- The ADS π and the probability measure P can be interpreted as random variables with realizations $\pi_i = \pi(\omega_i)$ and $P(\omega_i)$.

Proposition 2.3 (Pricing with a Deflator)
In an arbitrage-free market, the asset prices are given by

$$S_j = \mathrm{E}_P[H \cdot X_j].$$

Proof.

$$S_j = \sum_{i=1}^{I} \pi_i \cdot X_{ij} = \sum_{i=1}^{I} P(\omega_i) \frac{\pi(\omega_i)}{P(\omega_i)} X_j(\omega_i) = \mathrm{E}_P\left[\frac{\pi}{P} X_j\right] = \mathrm{E}_P[H \cdot X_j]$$

□

Remarks.

- The deflator H is the appropriate discount factor for payoffs at time $t = 1$.
- The deflator H ist a random variable with realizations $\pi(\omega_i)/P(\omega_i)$.
- On average, one needs to discount with the riskless rate r because

$$\mathrm{E}_P[H] = \sum_{i=1}^{I} P(\omega_i) \cdot \frac{\pi(\omega_i)}{P(\omega_i)} = \sum_{i=1}^{I} \pi(\omega_i) = \frac{1}{1+r}.$$

Interpretation of a Deflator: In a simple one-period equilibrium model with time additive utility, one can show that the deflator essentially equals the marginal rate of substitution between consumption at time $t = 0$ and $t = 1$ of the representative investor, i.e.[3]

$$H = \frac{\partial u/\partial c_1}{\partial u/\partial c_0} e^{-\delta},$$

where $u = u(c_0, c_1)$ is the utility function of the representative investor and δ is his time preference rate.

2.1.4 Risk-neutral Measure

Two important observations:

(i) If an investor buys all ADSs, he gets one euro at time $t = 1$ for sure, i.e. we get

$$\sum_{i=1}^{I} \pi_i = \frac{1}{1+r} \qquad \Rightarrow \qquad \sum_{i=1}^{I} \pi_i \cdot (1+r) = 1.$$

(ii) In an arbitrage-free market, it holds that $\pi_i > 0$, $i = 1, \dots, I$.

Proposition 2.4 (Risk-neutral Measure)
In an arbitrage-free market, there exists a probability measure Q defined via

$$Q(\omega_i) = q_i := \pi_i \cdot (1+r)$$

This measure is said to be the risk-neutral measure or equivalent martingale measure.

Proof. follows from (i) and (ii). □

Using the risk-neutral measure, there is a third representation of an asset price:

[3]See, e.g., Cochrane (2001, chapter 1) or Munk (2013, p. 194).

Proposition 2.5 (Risk-neutral Pricing)
In an arbitrage-free market, asset prices possess the following representation

$$S_j = \mathrm{E}_Q\Big[\frac{X_j}{R}\Big],$$

where $R := 1 + r$.

Proof.

$$S_j = \sum_{i=1}^{I} \pi_i \cdot X_{ij} = \sum_{i=1}^{I} \pi_i \cdot R \cdot \frac{X_{ij}}{R} = \sum_{i=1}^{I} q_i \cdot \frac{X_{ij}}{R} = \mathrm{E}_Q\Big[\frac{X_j}{R}\Big].$$

□

Remarks.

- Warning: The risk-neutral measure Q is different from the physical measure P.
- Roughly speaking, risk-neutral probabilities are compounded Arrow-Debreu prices.
- The notion “risk-neutral measure” comes from the fact that under this measure prices are discounted expected values where one needs to discount using the *riskless* interest rate.

To motivate the notion “equivalent martingale measure”, let us recall some basic definitions:

Definition 2.7 (Equivalent Probability Measures)
A Probability Measure Q is said to be continuous w.r.t. another probability measure P if for any event $A \subset \Omega$

$$P(A) = 0 \Longrightarrow Q(A) = 0.$$

If Q is continuous w.r.t. P and vice versa, then P and Q are said to be equivalent, i.e. both measures possess the same null sets.

Theorem 2.6 (Radon-Nikodym)
If Q is continuous w.r.t. P, then Q possesses a density w.r.t. Q, i.e. there exists a positive random variable $D = \frac{dQ}{dP}$ such that

$$\mathrm{E}_Q[Y] = \mathrm{E}_P[D \cdot Y]$$

for any random variable Y. This density is uniquely determined (P-a.s.).

Definition 2.8 (Stochastic Process in a Single-period Model)
Consider a single-period model with time points $t = 0, 1$. The finite sequence $\{Y(t)\}_{t=0,1}$ forms a stochastic process.

Definition 2.9 (Martingale in a Single-period Model)
Consider a single-period model with finite state space. Uncertainty at time $t = 1$ is modeled by some probability measure Q. A real-valued stochastic process is said to be a martingale w.r.t. Q (short: Q-martingale) if

$$\mathrm{E}_Q[Y(1)] = Y(0),$$

where it is assumed that $\mathrm{E}_Q[|Y(1)|] < \infty$.

We can now characterize risk-neutral measures as follows:

Proposition 2.7 (Properties of a Risk-neutral Measure)
(i) The discounted price processes $\{S_j, X_j/R\}$ are martingales under a risk-neutral measure Q.

(ii) The physical measure P and a risk-neutral measure Q are equivalent.

Proof. (i) follows from Proposition 2.5.

(ii) Since ADPs are strictly positive, we have $Q(\omega_i) > 0$ and $P(\omega_i) > 0$ which gives the desired result. □

Remarks.
- Proposition 2.7 is the reason why Q is said to be an **equivalent martingale measure**.
- From Propositions 2.3 and 2.5, we get

$$\mathrm{E}_Q\left[\frac{X_j}{R}\right] = S_j = \mathrm{E}_P[H \cdot X_j].$$

  Therefore, we obtain

$$\mathrm{E}_Q[X_j] = \mathrm{E}_P[\underbrace{H \cdot R}_{=dQ/dP} \cdot X_j],$$

  i.e. $H = (dQ/dP)/R$. Analogously, one can show $dP/dQ = 1/(H \cdot R)$.

2.1.5 Summary

Theorem 2.8 (No Arbitrage)
Consider a financial market with price vector S and payoff matrix X. If the assumptions (a)-(d) are satisfied, then the following statements are equivalent:

(i) The financial market contains no arbitrage opportunities (assumption (e)).

(ii) There exist strictly positive Arrow-Debreu prices π_i, $i = 1, \ldots, I$, such that

$$S_j = \sum_{i=1}^{I} \pi_i \cdot X_{ij}.$$

(iii) There exists a state-dependent discount factor H, the so-called deflator, such that for all $j = 0, \ldots, J$

$$S_j = \mathrm{E}_P[H \cdot X_j].$$

(iv) There exists an equivalent martingale measure Q such that the discounted price processes $\{S_j, X_j/R\}$ are martingales under Q, i.e. for all $j = 0, \ldots, J$

$$S_j = \mathrm{E}_Q\left[\frac{X_j}{R}\right]$$

Theorem 2.9 (Completeness)
Consider a financial market with price vector S and payoff matrix X. If the assumptions (a)-(e) are satisfied, then the following statements are equivalent:

(i) The financial market is complete, i.e. each additional asset can be replicated.

(ii) The Arrow-Debreu prices π_i, $i = 1, \ldots, I$, are uniquely determined.

(iii) There exists only one deflator H.

(iv) There exists only one equivalent martingale Q.

Remarks.
- In a complete market, the price of any additional asset is given by the price of its replicating portfolio.
- The value of this portfolio can be calculated applying the following theorem.

Theorem 2.10 (Pricing of Contingent Claims)
Consider an arbitrage-free and complete market with price vector S and payoff matrix X. Let $C \in I\!R^I$ be the payoff of a contingent claim. Its today's price S_C can be calculated by using one of the following formulas:

(i) $S_C = C'\pi = \sum_{i=1}^{I} \pi_i \cdot C(\omega_i)$ *(Pricing with ADSs).*

(ii) $S_C = \mathrm{E}_P[H \cdot C]$ *(Pricing with the deflator).*

(iii) $S_C = \mathrm{E}_Q\left[\frac{C}{R}\right]$ *(Pricing with the equivalent martingale measure).*

Proof. In an arbitrage-free and complete market, there exist ADPs which are uniquely determined. The no-arbitrage requirement together with the definition of an ADP gives formula (i). Formulas (ii) and (iii) follow from (i) using $H = \pi/P$ and $Q(\omega_i) = \pi_i \cdot R$. □

2.2 Multi-period Model

In the previous section we have modeled only two points in time:

$\boldsymbol{t = 0}$: The investor chooses the trading strategy.

$\boldsymbol{t = T}$: The investor liquidates his portfolio and gets the payoffs.

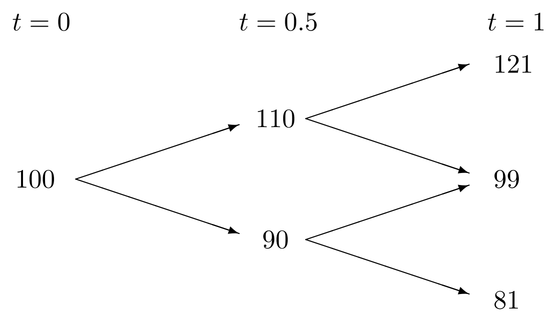
This has the following consequences:
- We have disregarded the *information flow* between 0 and T.
- Every trading strategy was a *static* strategy (”‘buy and hold”’), i.e. intermediate portfolio adjustments was not allowed.

In this section
- we will model the information flow between 0 and T.
  $\rightarrow$ Filtration
- and the investor can adjust his portfolio according to the arrival of new information about prices.
  $\rightarrow$ dynamic (self-financing) trading strategy.

2.2.1 Modeling Information Flow

Motivation

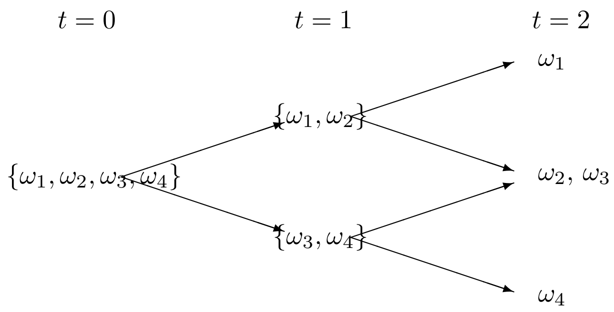
Consider the dynamics of a stock in a two-period binomial model:



- Due to the additional trading date at time $t = 0.5$, the investor gets more information about the stock price at time $t = 1$.
- At $t = 0$ it is possible that the stock price at time $t = 1$ equals 81, 99 or 121.
- If, however, the investor already knows that at time $t = 0.5$ the stock price equals 110 (90), then at time $t = 1$ the stock price can only take the values 121 or 99 (99 or 81).
- It is one of the goals of this section to model this additional piece of information!

Reference Example

To demonstrate how one can model the information flow, we consider the following example:



The *state space* thus equals: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$.

Example for the interpretation of the states.

ω_1 : S&P 500 stands at 2500 at $t = 1$ and at 3000 at time $t = 2$.
ω_2 : S&P 500 stands at 2500 at $t = 1$ and at 2000 at time $t = 2$.
ω_3 : S&P 500 stands at 1500 at $t = 1$ and at 2000 at time $t = 2$.
ω_4 : S&P 500 stands at 1500 at $t = 1$ and at 1000 at time $t = 2$.

First Approach to Model the Flow of Information

Definition 2.10 (Partition)
A partition of the state space is a set $\{A_1, \ldots, A_K\}$ of subsets $A_k \subset \Omega$ of a state space Ω such that

- *the subsets A_k are disjoint, i.e. $A_k \cap A_l = \emptyset$,*
- $\Omega = A_1 \cup \ldots \cup A_K$.

Example. $\{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}$.

Definition 2.11 (Information Structure)
An information structure is a sequence $\{F_t\}_t$ of partitions F_0, …, F_T such that

- $F_0 = \{\omega_1, \ldots, \omega_I\}$ *(”‘ At the beginning each state can occur.”’),*
- $F_T = \{\{\omega_1\}, \ldots, \{\omega_I\}\}$ *(”‘ At the end, one knows for sure which state has occurred. ”’),*
- *the partitions become finer over time.*

Example.

$F_0 = \{\{\omega_1, \omega_2, \omega_3, \omega_4\}\}$.

$F_1 = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}$.

$F_2 = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}$.

Remarks.

- **Problem:** This intuitive definition cannot be generalized to the continuous-time case.
- However, it provides a good intuition of what information flow actually means.

Second Approach to Model the Flow of Information: Filtration

Definition 2.12 (σ-Field)
A system $\mathcal{F}$ of subsets of the state space Ω is said to be a σ-field if

(i) $\Omega \in \mathcal{F}$,

(ii) $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$,

(iii) for every sequence $\{A_n\}_{n\in\mathbb{N}}$ *with* $A_n \in \mathcal{F}$ *we have* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Remark. For a finite σ-field, it is sufficient if finite (instead of infinite) unions of events belong to the σ-field (see requirement (iii) of Definition 2.12).

Example. $\{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\}$.

Definition 2.13 (Filtration)
Consider a state space $\mathcal{F}$. A family $\{\mathcal{F}_t\}_t$ of sub-σ-fields $\mathcal{F}_t$, $t \in \mathcal{I}$, is said to be a filtration if $\mathcal{F}_s \subset \mathcal{F}_t$, $s \leq t$ and $s, t \in \mathcal{I}$. Here $\mathcal{I}$ denotes some ordered index set.

Remark. In our discrete time model, we can always use filtrations with the following properties ($\mathcal{P}(\Omega)$ denotes the power-set of Ω):
- $\mathcal{F}_0 = \{\emptyset, \Omega\}$ ("'At the beginning, one has no information."'),
- $\mathcal{F}_T = \mathcal{P}(\Omega)$ ("'At the end one is fully informed."'),
- $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_T$ ("'The state of information increases over time."').

Example. $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_1 = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\}$, $\mathcal{F}_2 = \mathcal{P}(\Omega)$.

Some facts about filtrations:
- Filtrations are used to model the flow of information over time.
- At time t, one can decide if the event $A \in \mathcal{F}_t$ has occurred or not.
- Taking conditional expectation $\mathrm{E}[Y|\mathcal{F}_t]$ of a random variable Y means that one takes expectation on the basis of all information available at time t.

2.2.2 Description of the Multi-period Model

Properties of the Model:
- Trading takes place at $T+1$ dates, $t = 0, \ldots, T$.
- Investors can trade in $J+1$ assets with price processes $(S_0(t), \ldots, S_J(t))$, $t = 0, \ldots, T$.
- $S_j(t)$ denotes the price of the j-th asset at time t.
- The range of $S_j(t)$ is finite.
- Trading of an investor is modeled via his *trading strategy* $\varphi(t) = (\varphi_0(t), \ldots, \varphi_J(t))'$, where φ_j(t) denotes the number of asset j at time t in the investor's portfolio.
- A trading strategy needs to be *predictable*, i.e. on the basis of all information about prices $S(t-1)$ at time $t-1$ the investor selects his portfolio which he holds until time t.

- The value of a trading strategy at time t is given by:

$$V_\varphi(t) = \begin{cases} S(t)'\varphi(t) = \sum_{j=0}^{J} S_j(t)\varphi_j(t) & \text{für} \quad t = 1, \ldots, T \\ S(0)'\varphi(1) = \sum_{j=0}^{J} S_j(0)\varphi_j(1) & \text{für} \quad t = 0 \end{cases}$$

Definition 2.14 (Contingent Claim)
A contingent claim leads to a state-dependent payoff $C(t)$ at time t, which is measurable w.r.t. $\mathcal{F}_t$.

Remarks.
- A typical contingent claim is a European call with payoff $C(T) = \max\{S(T) - K; 0\}$, where $S(T)$ denotes the value of the underlying at time T and K the strike price.
- American options are no contingent claims in the sense of Definition 2.14, but one can generalize this definition accordingly.
- Recall the following **paradigm**: Arbitrage-free pricing of a contingent claim means finding a hedging (replication) strategy of the corresponding payoff.
- In the single-period model: The price of the replication strategy is then the price of the contingent claim.
- In the multi-period model we need an additional requirement: The replication strategy should not require intermediate inflows or outflows of cash.
- Only in this case, the initial costs are equal to the fair price of the contingent claim.
- This leads to the notion of a self-financing trading strategy.

Definition 2.15 (Self-financing Trading Strategy)
A trading strategy $\varphi(t) = (\varphi_1(t), \ldots, \varphi_J(t))$ is said to be self-financing if for $t = 1, \ldots, T-1$

$$S(t)'\varphi(t) = S(t)'\varphi(t+1) \qquad \Longleftrightarrow \qquad \sum_{j=0}^{J} S_j(t) \cdot \varphi_j(t) = \sum_{j=0}^{J} S_j(t) \cdot \varphi_j(t+1).$$

Remarks.
- To buy assets, the investor has to sell other ones.
- For this reason, the portfolio value changes only if prices change and not if the numbers of assets in the portfolio change.

Notation:
- Change of the value of strategy φ at time t:

$$\Delta V_\varphi(t) := V_\varphi(t) - V_\varphi(t-1)$$

- Price changes at time t

$$\Delta S(t) = (\Delta S_0(t), \ldots, \Delta S_J(t)) := (S_0(t) - S_0(t-1), \ldots, S_J(t) - S_J(t-1))$$

- Gains process of the trading strategy φ at time t:

$$\varphi(t)'\Delta S(t) = \sum_{j=0}^{J} \varphi_j(t)\Delta S_j(t)$$

- Cumulative gains process of the trading strategy φ at time t:

$$G_\varphi(t) := \sum_{\tau=1}^{t} \varphi(\tau)'\Delta S(\tau) = \sum_{\tau=1}^{t}\sum_{j=0}^{J} \varphi_j(\tau)\Delta S_j(\tau)$$

Proposition 2.11 (Characterization of a Self-financing Trading Strategy) *A self-financing trading strategy $\varphi(t)$ has the following properties:*

(i) The value of φ changes only due to price changes, i.e.

$$\Delta V_\varphi(t) = \varphi(t)'\Delta S(t).$$

(ii) The value of φ is equal to its initial value plus the cumulative gains, i.e.

$$V_\varphi(t) = V_\varphi(0) + G_\varphi(t) = V_\varphi(0) + \sum_{\tau=1}^{t} \varphi(\tau)'\Delta S(\tau).$$

Proof. (i) For a self-financing strategy φ, we obtain

$$\begin{aligned}
\Delta V_\varphi(t) &= V_\varphi(t) - V_\varphi(t-1) \\
&= \sum_{j=0}^{J} S_j(t)\varphi_j(t) - \sum_{j=0}^{J} S_j(t-1)\varphi_j(t-1) \\
&\overset{(*)}{=} \sum_{j=0}^{J} S_j(t)\varphi_j(t) - \sum_{j=0}^{J} S_j(t-1)\varphi_j(t) \\
&= \sum_{j=0}^{J} \varphi_j(t)\Big(S_j(t) - S_j(t-1)\Big) \\
&= \sum_{j=0}^{J} \varphi_j(t)\Delta S_j(t) \\
&= \varphi(t)'\Delta S(t).
\end{aligned}$$

The equality $(*)$ is valid because φ is self-financing.

(ii) Any (not necessary self-financing) trading strategy satisfies

$$V_\varphi(t) = V_\varphi(0) + \sum_{\tau=1}^{t} \Delta V_\varphi(\tau).$$

Since φ is assumed to be self-financing, by (i), we obtain

$$\begin{aligned}
V_\varphi(t) &= V_\varphi(0) + \sum_{\tau=1}^{t} \Delta V_\varphi(\tau). \\
&\overset{(i)}{=} V_\varphi(0) + \sum_{\tau=1}^{t} \varphi(\tau)'\Delta S(\tau) \\
&= V_\varphi(0) + G_\varphi(t).
\end{aligned}$$

□

Remarks.

- In continuous-time models, the relation (i) is still valid if Δ is replace by d.
- Unfortunately, the intuitive relation

$$S(t)'\varphi(t) = S(t)'\varphi(t+1)$$

  has no continuous-time equivalent.

2.2.3 Some Basic Definitions Revisited

The notion of a self-financing trading strategy is also important to generalize the following definitions to the multi-period setting:

Definition 2.16 (Arbitrage)
We distinguish two kinds of arbitrage opportunities:

(i) A free lunch is a self-financing trading strategy φ with

$$V_\varphi(0) < 0 \quad \textit{und} \quad V_\varphi(T) \geq 0.$$

(ii) A free lottery is a self-financing trading strategy φ with

$$V_\varphi(0) = 0 \quad \textit{und} \quad V_\varphi(T) \geq 0 \quad \textit{und} \quad V_\varphi(T) \neq 0.$$

Definition 2.17 (Replication of a Contingent Claim)
A contingent claim with payoff $C(T)$ is said to be replicable if there exists a self-financing trading strategy φ such that

$$C(T) = V_\varphi(T).$$

The trading strategy φ is said to be a replication (hedging) strategy for the claim $C(T)$.

Remarks.

- By definition, a replication strategy of a claim leads to the same payoff as the claim.
- In general, replication strategies are *dynamic* strategies, i.e. at every trading date the hedge portfolio needs to be adjusted according to the arrival of new information about prices.

The definition of a complete market remains the same:

Definition 2.18 (Completeness)
An arbitrage-free market is said to be complete if each contingent claim can be replicated.

2.2.4 Assumptions

(a) All investors agree upon which states cannot occur at times $t > 0$.

(b) All investors agree upon the range of the price processes $S(t)$.

(c) At time $t = 0$ there is only one market price for each security and each investor can buy and sell as many securities as desired without changing the market price (price taker).

(d) Frictionless markets, i.e.

- assets are divisible,
- no restrictions on short sales,[4]
- no transaction costs,
- no taxes.

(e) The asset market contains no arbitrage opportunities, i.e. there are neither free lunches nor free lotteries.

2.2.5 Main Results

Conventions:

- The 0-th asset is a money market account, i.e. it is locally risk-free with

$$S_0(0) = 1 \qquad \text{und} \qquad S_0(t) = \prod_{u=1}^{t} (1 + r_u), \quad t = 1, \ldots, T,$$

  where r_u denotes the interest rate for the period $[\tau - 1, \tau]$.
- The discounted (normalized) price processes are defined by

$$Z_j(t) := \frac{S_j(t)}{S_0(t)}.$$

- The money market account is thus used as a numeraire.

Definition 2.19 (Martingal)
A real-valued adapted process $\{Y(t)\}_t$, $t = 0, \ldots, T$, with $\mathrm{E}[|Y(t)|] < \infty$ is a martingale w.r.t. a probability measure Q if for $s \leq t$

$$\mathrm{E}_Q[Y(t)|\mathcal{F}_s] = Y(s).$$

Remark. A martingale is thus a process being on average stationary.

Important Facts:

[4] A short sale of an assets is equivalent to selling an asset one does not own. An investor can do this by borrowing the asset from a third party and selling the borrowed security. The net position is a cash inflow equal to the price of the security and a liability (borrowing) of the security to the third party. In abstract mathematical terms, a short sale corresponds to buying a negative amount of a security.

- A multi-period model consists of a sequence of single-period models.
- Therefore, a multi-period model is arbitrage-free if all single-period models are arbitrage-free (See, e.g., Bingham/Kiesel (2004), Chapter 4).

For these reasons, the following three theorems hold:

Theorem 2.12 (Characterization of No Arbitrage)
If assumptions (a)-(d) are satisfied, the following statements are equivalent:

(i) The financial market contains no arbitrage opportunities (assumption (e)).

(ii) There exist a state-dependent discount factor H, the so-called deflator, such that for all time points $s, t \in \{0, \dots, T\}$ with $s \le t$ and all assets $j = 0, \dots, J$ we have

$$S_j(s) = \frac{\mathrm{E}_P[H(t)S_j(t)|\mathcal{F}_s]}{H(s)},$$

i.e. all prices can be represented as discounted expectations under the physical measure.

(iii) There exists a risk-neutral probability measure Q such that all discounted price processes $Z_j(t)$ are martingales under Q, i.e. for all time points $s, t \in \{0, \dots, T\}$ with $s \le t$ and all assets $j = 1, \dots, J$ we have

$$Z_j(s) = \mathrm{E}_Q[Z_j(t)|\mathcal{F}_s] \qquad \Longleftrightarrow \qquad \frac{S_j(s)}{S_0(s)} = \mathrm{E}_Q\Big[\frac{S_j(t)}{S_0(t)}\Big|\mathcal{F}_s\Big]$$

Remarks.
- The equivalence “No Arbitrage $\Leftrightarrow$ Existence of Q” is known as **Fundamental Theorem of Asset Pricing**.
- Unfortunately, in continuous-time models this equivalence breaks down.
- However, the implication “Existence of $Q \Rightarrow$ No Arbitrage”’ is still valid,

Theorem 2.13 (Pricing of Contingent Claims)
Consider a complete financial market satisfying assumptions (a)-(e) and a contingent claim with payoff $C(T)$. Then we obtain for $s, t \in \{0, \dots, T\}$ with $s \le t$:

(i) The replication strategy φ for the claim is unique, i.e. the arbitrage-free price of the claim is uniquely determined by

$$C(s) = V_\varphi(s).$$

(ii) Using the deflator, we can calculate the claim’s time-s price

$$C(s) = \frac{\mathrm{E}_P[H(t)C(t)|\mathcal{F}_s]}{H(s)}.$$

(iii) The following risk-neutral pricing formula holds:

$$C(s) = S_0(s) \cdot \mathrm{E}_Q\Big[\frac{C(t)}{S_0(t)}\Big|\mathcal{F}_s\Big].$$

Remarks.

- Result (i) says that the today's price of a claim is uniquely determined by its replication costs, i.e. $C(0) = V_\varphi(0)$.
- From (iii) we get
$$\frac{C(s)}{S_0(s)} = \mathrm{E}_Q\Big[\frac{C(t)}{S_0(t)}\Big|\mathcal{F}_s\Big],$$
i.e. the discounted price processes of contingent claims are Q-martingales as well. Note that one needs to discount with the money market account.
- If interest rates are constant, then from (iii)
$$C(0) = \mathrm{E}_Q\Big[\frac{C(T)}{(1+r)^T}\Big] = \frac{\mathrm{E}_Q[C(T)]}{(1+r)^T}.$$
- Pricing of contingent claims boils down to computing expectations under the risk-neutral measure.

Theorem 2.14 (Completeness)
Consider a financial market satisfying assumptions (a)-(e). Then the following statements are equivalent

(i) The market is complete.

(ii) There exists a unique deflator H.

(iii) There exists a unique martingale measure Q.

2.2.6 Summary

The following results are important:

- It is always valid that a financial market contains no arbitrage opportunities if an equivalent martingale measure exists.
- In our discrete model, the converse is also valid: If the market contains no arbitrage opportunities, there exists a martingale measure.
- If a martingale measure Q exists, the discounted price processes are Q-martingales.
- If a unique martingale measure exists, the market is arbitrage-free and complete.
- If a unique martingale measure exists, the today's price of a contingent claim is given by the following Q-expectation
$$C(0) = \mathrm{E}_Q\Big[\frac{C(T)}{S_0(T)}\Big].$$

3 Introduction to Stochastic Calculus

3.1 Motivation

3.1.1 Why Stochastic Calculus?

- In the previous section, the investor was able to trade at finitely many time points $t = 0, \dots, T$ only.
- In a continuous-time model, trading takes place continuously, i.e. at any point in the time span $[0, T]$.
- Consequently, price dynamics are modeled continuously.
- In continuous-time models asset dynamics are modeled via stochastic differential equations.

Example "Black-Scholes Model". Two investment opportunities:

- money market account: riskless asset with constant interest rate r modeled via an ordinary differential equation (ODE):

$$dM(t) = M(t)rdt, \qquad M(0) = 1.$$

- stock: risky asset modeled via a stochastic differential equation (SDE)

$$dS(t) = S(t)\Big[\mu dt + \sigma dW(t)\Big], \qquad S(0) = s_0,$$

  where W is a Brownian motion.

In this section we answer the following questions:

Q1 What is a Brownian motion?

Q2 What is the meaning of $dW(t)$? $\rightarrow$ Stochastic Integral

Q3 What are stochastic differential equations and do explicit solutions exist? $\rightarrow$ Variation of Constant

Q4 How does the dynamics of contingent claims look like? $\rightarrow$ Ito Formula

3.1.2 What is a Reasonable Model for Stock Dynamics?

Model for the Money Market Account

The solution of the ODE $B'(t) = r \cdot B(t)$ with $B(0) = 1$ is

$$B(t) = b_0 \exp(rt)$$

and thus

$$\ln(B(t)) = \ln(b_0) + r \cdot t.$$

Model for the Stock

The result for the money market account suggests the following model for the stock dynamics:

$$\ln(S(t)) = \ln(s_0) + \alpha \cdot t + \text{“randomness”}$$

Notation: $Y(t) =$ “randomness”, i.e. $Y(t) = \ln(S(t)) - \ln(s_0) - \alpha \cdot t$

Reasonable properties of Y:

1) no tendency, i.e. $\mathrm{E}[Y(t)] = 0$,
2) increasing with time t, i.e. $\mathrm{Var}(Y(t))$ increases with t,
3) no memory, i.e. for $Y(t) = Y(s) + [Y(t) - Y(s)]$ the difference $[Y(t) - Y(s)]$
   - depends only on $t - s$ and
   - is independent of $Y(s)$,

Properties which would make life easier:

4) normally distributed, i.e. $Y(t) \sim \mathcal{N}(0, \sigma^2 t)$, $\sigma > 0$,
5) continuous.

3.2 On Stochastic Processes

Definition 3.1 (Stochastic Process)
A stochastic process X defined on $(\Omega, \mathcal{F})$ is a collection of random variables $\{X_t)\}_{t \in \mathcal{I}}$, i.e. $\sigma\{X_t\} \subset \mathcal{F}$.

Remark. As otherwise stated, our stochastic processes take values in the measurable space $(I\!R^n, \mathcal{B}(I\!R^n))$, i.e. the n-dimensional Euclidian space equipped with σ-field of Borel sets. Therefore, every X_t is $(\mathcal{F}, \mathcal{B}(I\!R^n))$-measurable.

3.2.1 Cadlag and Caglad

Definition 3.2 (Cadlag and Caglad Functions)
Let $f : [0, T] \to I\!R^n$ be a function with existing left and right limits, i.e. for each $t \in [0, T]$ the limits

$$f(t-) = \lim_{s \nearrow t} f(s), \qquad f(t+) = \lim_{s \searrow t} f(s)$$

exist. The function f is cadlag if $f(t) = f(t+)$ and caglad if $f(t) = f(t-)$.

A cadlag function cannot jump around too widly as the next proposition shows.

Proposition 3.1
(i) A cadlag function f can have at most a countable number of discontinuities, i.e. $\{t \in [0,T] : f(t) \neq f(t-)\}$ is countable.

(ii) For any $\varepsilon > 0$, the number of discontinuities on $[0,T]$ larger than ε is finite.

Proof. See Fristedt/Gray (1997).

3.2.2 Modification, Indistinguishability, Measurability

Recall the following definition:

Definition 3.3 (Filtration)
Consider a σ-field $\mathcal{F}$. A family $\{\mathcal{F}_t\}$ of sub-σ-fields $\mathcal{F}_t$, $t \in \mathcal{I}$, is said to be a filtration if $\mathcal{F}_s \subset \mathcal{F}_t$, $s \leq t$ and $s, t \in \mathcal{I}$. Here $\mathcal{I}$ denotes some ordered index set.

Definition 3.4 (Adapted Process)
A process X is said to be adapted to the filtration $\{\mathcal{F}_t\}_{t\in\mathcal{I}}$ if every X_t is measurable w.r.t. $\mathcal{F}_t$, i.e. $\sigma\{X_t\} \subset \mathcal{F}_t$.

Remarks.
- The notation $\{X_t, \mathcal{F}_t\}_{t\in\mathcal{I}}$ indicates that the process X is adapted to $\{\mathcal{F}_t\}_{t\in\mathcal{I}}$.
- For fixed $\omega \in \Omega$, the realization $\{X_t(\omega)\}_t$ is said to be a path of X. For this reason, a stochastic process can be interpreted as function-valued random variable.

Definition 3.5 (Modification, Indistinguishability)
Let X and Y be stochastic processes.

(i) Y is a modification of X if

$$P(\omega : X_t(\omega) = Y_t(\omega)) = 1 \qquad \text{for all } t \geq 0.$$

(ii) X and Y are indistinguishable if

$$P(\omega : X_t(\omega) = Y_t(\omega) \text{ for all } t \geq 0) = 1,$$

i.e. X and Y have the same paths P-a.s.

Example. Consider a positive random variable τ with a continuous distribution. Set $X_t \equiv 0$ and $Y_t = \mathbf{1}_{\{t=\tau\}}$. Then Y is a modification of X, but $P(X_t = Y_t \text{ for all } t \geq 0) = 0$.

Proposition 3.2 (Modification, Indistinguishability)
Let X and Y be stochastic processes.

(i) If X and Y are indistinguishable, then Y is a modification of X and vice versa.

(ii) If X and Y have right-continuous paths and Y is a modification of X, then X and Y are indistinguishable.

(iii) Let X be adapted to $\{\mathcal{F}_t\}$. If Y is a modification of X and $\mathcal{F}_0$ contains all P-null sets of $\mathcal{F}$, then Y is adapted to $\{\mathcal{F}_t\}_t$.

Proof. (i) is obvious.

(ii) Since Y is a modification of X, we have

$$P(\omega | X_t(\omega) = Y_t(\omega) \text{ for all } t \geq [0,\infty) \cap \mathbb{Q}) = 1.$$

By the right-continuity assumption, the claim follows.

(iii) Exercise. □

If X is an (adapted) stochastic process, then every X_t is $(\mathcal{F}, \mathcal{B}(\mathbb{R}^n))$-measurable $((\mathcal{F}_t, \mathcal{B}(\mathbb{R}^n))$-measurable). However, X is actually a function-valued random variable, and so, for technical reasons, it is often convenient to have some joint measurability properties.

Definition 3.6 (Measurable Stochastic Process)
A stochastic process is said to measurable, if for each $A \in \mathcal{B}(\mathbb{R}^n)$ the set $\{(t,\omega) : X_t(\omega) \in A\}$ belongs to the product σ-field $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{F}$, i.e. if the mapping

$$(t,\omega) \mapsto X_t(\omega) : ([0,\infty) \times \Omega, \mathcal{B}([0,\infty)) \otimes \mathcal{F}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$$

is measurable.

Definition 3.7 (Progressively Measurable Stochastic Process)
A stochastic process is said to progressively measurable w.r.t. the filtration $\{\mathcal{F}_t\}$, if for each $t \geq 0$ and $A \in \mathcal{B}(\mathbb{R}^n)$ the set $\{(s,\omega) \in [0,t] \times \Omega : X_s(\omega) \in A\}$ belongs to the product σ-field $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$, i.e. if the mapping

$$(s,\omega) \mapsto X_s(\omega) : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$$

is measurable for each $t \geq 0$.

Remark. Any progressively measurable process is measurable and adapted (Exercise). The following proposition provides the extent to which the converse is true.

Proposition 3.3
If a stochastic process is measurable and adapted to the filtration $\{\mathcal{F}_t\}$, then it has a modification which is progressively measurable w.r.t. $\{\mathcal{F}_t\}$.

Proof. See Karatzas/Shreve (1991, p. 5).

Fortunately, if a stochastic process is left- or right-continuous, a stronger result can be proved:

Proposition 3.4
If the stochastic process X is adapted to the filtration $\{\mathcal{F}_t\}$ and right- or left-continuous, then X is also progressively measurable w.r.t. $\{\mathcal{F}_t\}$.

Proof. See Karatzas/Shreve (1991, p. 5).

Remark. If X is right- or left-continuous, but not necessarily adapted to $\{\mathcal{F}_t\}$, then X is at least measurable.

3.2.3 Stopping Times

Consider a measurable space $(\Omega, \mathcal{F})$. An $\mathcal{F}$- measurable random variable $\tau : \Omega \to \mathcal{I} \cup \{\infty\}$ where $\mathcal{I} = [0, \infty)$ or $\mathcal{I} = I\!N$ is said to be a **random time**. Two special classes of random times are of particular importance.

Definition 3.8 (Stopping Time, Optional Time)
(i) A random time is a stopping time w.r.t. the filtration $\{\mathcal{F}_t\}_{t\in\mathcal{I}}$ if the event $\{\tau \leq t\} \in \mathcal{F}_t$ for every $t \in \mathcal{I}$.

(ii) A random time is an optional time w.r.t. the filtration $\{\mathcal{F}_t\}_{t\in\mathcal{I}}$ if the event $\{\tau < t\} \in \mathcal{F}_t$ for every $t \in \mathcal{I}$.

Propositions 3.5 and 3.6 show how both concepts are connected.

Proposition 3.5
If τ is optional and $c > 0$ is a positive constant, then $\tau + c$ is a stopping time.

Proof. Exercise.

Definition 3.9 (Continuity of Filtrations)
(i) A filtration $\{\mathcal{G}_t\}$ is said to be right-continuous if

$$\mathcal{G}_t = \mathcal{G}_{t+} := \bigcap_{\varepsilon>0} \mathcal{G}_{t+\varepsilon}$$

(ii) A filtration $\{\mathcal{G}_t\}$ is said to be left-continuous if

$$\mathcal{G}_t = \mathcal{G}_{t-} := \sigma\left\{\bigcup_{s<t} \mathcal{G}_s\right\}$$

Remark. Let X be a stochastic process.

- Loosely speaking, left-continuity of $\{F_t^X\}$ means that X_t can be guessed by observing X_s, $0 \leq s < t$.
- Right-continuity means intuitively that if X_s has been observed for $0 \leq s \leq t$, then one cannot learn more by peeking infinitesimally far into the future.

Proposition 3.6
Every stopping time is optional, and both concepts coincide if the filtration is right-continuous.

Proof. Consider a stopping time τ. Then $\{\tau < t\} = \bigcup_{t>\varepsilon>0}\{T \leq t - \varepsilon\}$ and $\{T \leq t-\varepsilon\} \in \mathcal{F}_{t-\varepsilon} \subset \mathcal{F}_t$, i.e. τ is optional. On the other hand, consider an optional time τ w.r.t. a right-continuous filtration $\{\mathcal{F}_t\}$. Since $\{\tau \leq t\} = \bigcap_{t+\varepsilon>u>t}\{T < u\}$, we have $\{\tau \leq t\} \in \mathcal{F}_{t+\varepsilon}$ for every $\varepsilon > 0$ implying $\{\tau \leq t\} = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_{t+} = \mathcal{F}_t$. □

Remark. In our applications, we will always work with right-continuous filtrations.[5]

[5] See Definition 3.18.

Proposition 3.7
(i) Let τ_1 and τ_2 be stopping times. Then $\tau_1 \wedge \tau_2 = \min\{\tau_1, \tau_2\}$, $\tau_1 \vee \tau_2 = \max\{\tau_1, \tau_2\}$, $\tau_1 + \tau_2$, and $\alpha\tau_1$ with $\alpha \geq 1$ are stopping times.

(ii) Let $\{\tau_n\}_{n\in\mathbb{N}}$ be a sequence of optional times. Then the random times

$$\sup_{n\geq 1} \tau_n, \quad \inf_{n\geq 1} \tau_n, \quad \overline{\lim}_{n\to\infty}\tau_n, \quad \underline{\lim}_{n\to\infty}\tau_n$$

are all optional. Furthermore, if the τ_n's are stopping times, then so is $\sup_{n\geq 1} \tau_n$.

Proof. Exercise.

Proposition and Definition 3.8 (Stopping Time σ-Field)
Let τ be a stopping time of the filtration $\{\mathcal{F}_t\}$. Then the set

$$\{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\}$$

is a σ-field, the so-called stopping time σ-field $\mathcal{F}_\tau$.

Proof. Exercise.

Proposition 3.9 (Properties of a Stopped σ-Field)
Let τ_1 and τ_2 be stopping times.

(i) If $\tau_1 = t \in [0, \infty)$, then $\mathcal{F}_{\tau_1} = \mathcal{F}_t$.

(ii) If $\tau_1 \leq \tau_2$, then $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$.

(iii) $\mathcal{F}_{\tau_1\wedge\tau_2} = \mathcal{F}_{\tau_1} \cap \mathcal{F}_{\tau_2}$ and the events

$$\{\tau_1 < \tau_2\}, \quad \{\tau_2 < \tau_1\}, \quad \{\tau_1 \leq \tau_2\}, \quad \{\tau_2 \leq \tau_1\}, \quad \{\tau_1 = \tau_2\}$$

belong to $\mathcal{F}_{\tau_1} \cap \mathcal{F}_{\tau_2}$.

Proof. Exercise.

Definition 3.10 (Stopped Process)
(i) If X is a stochastic process and τ is random time, we define the function X_τ on the event $\{\tau < \infty\}$ by

$$X_\tau(\omega) := X_{\tau(\omega)}(\omega)$$

If $X_\infty(\omega)$ is defined for all $\omega \in \Omega$, then X_τ can also be defined on Ω, by setting $X_\tau(\omega) := X_\infty(\omega)$ on $\{\tau = \infty\}$.

(ii) The process stopped at time τ is the process $\{X_{t\wedge\tau}\}_{t\in[0,\infty)}$.

Proposition 3.10 (Measurability of a Stopped Process)
(i) If the process X is measurable and the random time τ is finite, then the function X_τ is a random variable.

(ii) If the process X is progressively measurable w.r.t. $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ and τ is a stopping time of this filtration, then the random variable X_τ, defined on the set $\{\tau < \infty\} \in \mathcal{F}_\tau$, is $\mathcal{F}_\tau$-measurable, and the stopped process $\{X_{\tau\wedge t}, \mathcal{F}_t\}_{t\in[0,\infty)}$ is progressively measurable.

Proof. Karatzas/Shreve (1991, p. 5 and p. 9).

Remark. The following result is obvious: If X is adapted and cadlag and τ is a stopping time, then

$$X_{t\wedge\tau} = X_t \mathbf{1}_{\{t<\tau\}} + X_\tau \mathbf{1}_{\{t\geq\tau\}}.$$

is also adapted and cadlag.

Definition 3.11 (Hitting Time)
Let X be a stochastic process and let $A \in \mathcal{B}(\mathbb{R}^n)$ be a Borel set in $\mathbb{R}^n$. Define

$$\tau_A(\omega) = \inf\{t > 0 : X_t(\omega) \in A\}.$$

Then τ is called a hitting time of A for X.

Proposition 3.11
Let X be a right-continuous process adapted to the filtration $\{\mathcal{F}_t\}$.

(i) If A_o is an open set, then the hitting time of A_o is an optional time.

(ii) If A_c is a closed set, then the random variable

$$\tilde{\tau}_{A_c}(\omega) = \inf\{t > 0 : X_t(\omega) \in A_c \text{ or } X_{t-}(\omega) \in A_c\}.$$

is a stopping time.

Proof. (i) For simplicity $n = 1$. We have $\{\tau_{A_o} < t\} = \bigcup_{s\in\mathbb{Q}\cap[0,t)}\{X_s \in A_o\}$. The inclusion $\supset$ is obvious. Suppose that the inclusion $\subset$ does not hold. Then we have $X_i \in A_o$ for some $i \in \mathbb{R} \setminus \mathbb{Q}$ and $i < t$, but $X_q \notin A_o$ for all $q \in \mathbb{Q}$ with $q < t$. Since A_o is open, there exists a $\delta > 0$ such that $(X_i - \delta, X_i + \delta) \subset A_o$. Therefore, $X_q \notin (X_i - \delta, X_i + \delta)$. However, then there exists a sequence $(q_n)_{n\in\mathbb{N}}$ with $q_n \searrow s_0$ and $q_n \in \mathbb{Q} \cap [0, t)$, but $X_{q_n} \not\to X_i$ which is a contradiction to the right-continuity of X.

(ii) See Karatzas/Shreve (1991, p. 39) and Protter (2005, p. 5). □

Remarks.

- If the filtration is right-continuous, then the hitting time τ_{A_o} is a stopping time.
- If X is continuous , then the stopping time $\tilde{\tau}_{A_c}$ is a hitting time of A_c.
- Even if X is continuous, in general the hitting time of A_o is not a stopping time. For example, suppose that X is real-valued, that for some ω, $X_t(\omega) \leq 1$ for $t \leq 1$, and that $X_1(\omega) = 1$. Let $A_o = (1, \infty)$. Without looking slightly ahead of time 1, we cannot tell wheter or not $\tau_{A_o} = 1$.

3.3 Prominent Examples of Stochastic Processes

Theorem and Definition 3.12 (Brownian Motion)
(i) A Brownian motion is an adapted, real-valued process $\{W_t, \mathcal{F}_t\}_{t\in[0,\infty)}$ such that $W_0 = 0$ a.s. and

- *$W_t - W_s \sim \mathcal{N}(0, t-s)$ for $0 \le s < t$ "stationary increments"*
- *$W_t - W_s$ independent of $\mathcal{F}_s$ for $0 \le s < t$ "independent increments"*

exists and is said to be a one-dimensional Brownian motion.

(ii) An n-dimensional Brownian motion $W := (W_1, \dots, W_n)$ is an n-dimensional vector of independent one-dimensional Brownian motions.

Proof. Applying Kolmogorov's extension theorem,[6] it is straightforward to prove the existence of a Brownian motion. See, e.g., Karatzas/Shreve (1991, pp. 47ff). □

Theorem 3.13 (Kolmogorov's Continuity Theorem)
Suppose that the process $X = \{X_t\}_{t\in[0,\infty)}$ satisfies the following condition: For all $T > 0$ there exist positive constants α, β, C such that

$$\mathrm{E}[|X_t - X_s|^\alpha] \le C \cdot |t-s|^{1+\beta}, \qquad 0 \le s, t \le T.$$

Then there exists a continuous modification of X.

Proof. See Karatzas/Shreve (1991, pp.53ff).

Corollary 3.14 (Continuous Version of the Brownian Motion)
An n-dimensional Brownian motion possesses a continuous modification.

Proof. One can show that

$$\mathrm{E}[|W_t - W_s|^4] = n(n+2)|t-s|^2.$$

Hence, we can apply Kolmogorov's continuity theorem with $\alpha = 4$, $C = n(n+2)$, and $\beta = 1$. □

Remark. We will always work with this continuous version.

Definition 3.12 (Poisson Process)
A Poisson process with intensity $\lambda > 0$ is an adapted, integer-valued cadlag process $N = \{N_t, \mathcal{F}_t\}_{t[0,\infty)}$ such that $N_0 = 0$, and for $0 \le t < \infty$, the difference $N_t - N_s$ is independent of $\mathcal{F}_s$ and is Poisson distributed with mean $\lambda(t-s)$.[7]

3.4 Martingales

Definition 3.13 (Martingale, Supermartingale, Submartingale)
A process $X = \{X_t\}_{t\in[0,\infty)}$ adapted to the the filtration $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ is called a martingale (resp. supermartingale, submartingale) with respect to $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ if (i) $X_t \in L^1(dP)$, i.e. $\mathrm{E}[|X_t|] < \infty$.

[6] Kolmogorov's extension theorem basically says that for given (finite-dimensional) distributions satisfying a certain consistency condition there exists a probability space and a stochastic process such that its (finite-dimensional) distributions coincide with the given distributions. See, e.g., Karatzas/Shreve (1991, p. 50).

[7] A random variable Y taking values in $\mathbb{N} \cup \{0\}$ is Poisson distributed with parameter $\lambda > 0$ if $P(Y = n) = e^{-\lambda}\lambda^n/(n!)$, $n \in \mathbb{N}$.

(ii) if $s \leq t$, then $\mathrm{E}[X_t|\mathcal{F}_s] = X_s$, a.s. (resp. $\mathrm{E}[X_t|\mathcal{F}_s] \leq X_s$, resp. $\mathrm{E}[X_t|\mathcal{F}_s] \geq X_s$).

Examples.

- Let W be a one-dimensional Brownian motion. Then W is a martingale and W^2 is a submartingale.
- The arithmetic Brownian motion $X_t := \mu t + \sigma W_t$, μ, $\sigma \in I\!R$, is a martingale for $\mu = 0$, a submartingale for $\mu > 0$, and a supermartingale for $\mu < 0$.
- Let N be a Poisson process with intensity λ. Then $\{N_t - \lambda t\}_t$ is a martingale.

Definition 3.14 (Square-Integrable)
Let $X = \{X_t, \mathcal{F}_t\}_{t\in[0,\infty)}$ be a right-continuous martingale. We say that X is square-integrable if $\mathrm{E}[X_t^2] < \infty$ for every $t \geq 0$. If, in addition, $X_0 = 0$ a.s. we write $X \in \mathcal{M}_2$ (or $X \in \mathcal{M}_2^c$ if X is also continuous).

Proposition 3.15 (Functions of Submartingales)
(i) Let $X = (X_1, \ldots, X_n)$ with real-valued components X_k be a martingale w.r.t. $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ and let $\varphi : I\!R^n \to I\!R$ be a convex function such that $\mathrm{E}[|\varphi(X_t)|] < \infty$ for all $t \in [0,\infty)$. Then $\varphi(X)$ is a submartingale w.r.t. $\{\mathcal{F}_t\}_{t\in[0,\infty)}$

(ii) Let $X = (X_1, \ldots, X_n)$ with real-valued components X_k be a submartingale w.r.t. $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ and let $\varphi : I\!R^n \to I\!R$ be a convex, non-decreasing function such that $\mathrm{E}[|\varphi(X_t)|] < \infty$ for all $t \in [0,\infty)$. Then $\varphi(X)$ is a submartingale w.r.t. $\{\mathcal{F}_t\}_{t\in[0,\infty)}$

Proof. follows from Jensen's inequality (Exercise). □

For any $X \in \mathcal{M}_2$ and $0 \leq t < \infty$, we define

$$||X||_t := \sqrt{\mathrm{E}[X_t^2]} \quad \text{and} \quad ||X|| := \sum_{n=1}^{\infty} \frac{||X||_n \wedge 1}{2^n}.$$

Note that $||X - Y||$ defines a the pseudo-metric on $\mathcal{M}_2$. If we identify indistinguishable processes, the pseudo-metric $||X - Y||$ becomes a metric on $\mathcal{M}_2$.[8]

Proposition 3.16
If we identify indistinguishable processes, the pair $(\mathcal{M}_2, ||)$ is a complete metric space, and $\mathcal{M}_2^c$ is a closed subset of $\mathcal{M}_2$.

Proof. See, e.g., Karatzas/Shreve (1991, pp. 37f).

An important question now arises: What happens to the martingale property of a stopped martingale? For bounded stopping times, nothing changes.

Theorem 3.17 (Optional Sampling Theorem, 1st Version)
Let $X = \{X_t, \mathcal{F}_t\}_{t\in[0,\infty)}$ be a right-continuous submartingale (supermartingale) and let τ_1 and τ_2 be bounded stopping times of the filtration $\{\mathcal{F}_t\}$ with $\tau_1 \leq \tau_2 \leq c \in I\!R$ a.s. Then

$$\mathrm{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}] \geq (\leq) X_{\tau_1}$$

[8] This means: (i) $||X - Y|| \geq 0$ and $||X - Y|| = 0$ iff $X = Y$. (ii) $||X - Y|| = ||Y - X||$. (iii) $||X - Y|| = ||X - Z|| + ||Z - Y||$.

If the stopping times are unbounded, in general, this result does only hold if X can be extended to ∞ such that $\tilde{X} := \{X_t, \mathcal{F}_t\}_{t\in[0,\infty]}$ is still a sub- or supermartingale.

Theorem 3.18 (Sub- and Supermartingale Convergence)
Let $X = \{X_t, \mathcal{F}_t\}_{t\in[0,\infty)}$ be a right-continuous submartingale (supermartingale) and assume $\sup_{t\geq 0} \mathrm{E}[X_t^+] < \infty$ ($\sup_{t\geq 0} \mathrm{E}[X_t^-] < \infty$). Then the limit $X_\infty(\omega) = \lim_{t\to\infty} X_t(\omega)$ exists for a.e. $\omega \in \Omega$ and $\mathrm{E}|X_\infty| < \infty$.

Proof. See Karatzas/Shreve (1991, pp. 17f).

This proposition ensures that a limit exists, but it is not clear at all if $\tilde{X}$ is a sub- or supermartingale. To solve this problem, we need to require the following condition:

Definition 3.15 (Uniform Integrability)
A familiy of random variables $(U)_{a\in\mathcal{A}}$ is uniformly integrable (UI) if

$$\lim_{n\to\infty} \sup_a \int_{\{|U_\alpha|\geq n\}} |U_\alpha| dP = 0.$$

Using the definition, it is sometimes not easy to check if a family of random variables is uniformly integrable.

Theorem 3.19 (Characterization of UI)
Let $(U)_{a\in\mathcal{A}}$ be a subset of L^1. The following are equivalent:

(i) $(U)_{a\in\mathcal{A}}$ is uniformly integrable.

(ii) There exists a positive, increasing, convex function g defined on $[0,\infty)$ such that

$$\lim_{x\to\infty} \frac{g(x)}{x} = +\infty \quad \text{and} \quad \sup_{a\in\mathcal{A}} \mathrm{E}[g \circ |U_\alpha|] < \infty.$$

Remark. A good choice is often $g(x) = x^{1+\varepsilon}$ with $\varepsilon > 0$.

Proposition 3.20
The following conditions are equivalent for a nonnegative, right-continuous submartingale $X = \{X_t, \mathcal{F}_t\}_{t\in[0,\infty)}$:

(i) It is a uniformly integrable family of random variables.

(ii) It converges in L^1, as $t \to \infty$.

(iii) It converges P-a.s. ($t \to \infty$) to an integrable random variable X_∞, such that $\{X_t, \mathcal{F}_t\}_{t\in[0,\infty]}$ is a submartingale.

For the implication "(i)⇒(ii)⇒(iii)", the assumption of nonnegativity can be dropped.

Theorem 3.21 (UI Martingales can be closed)
The following conditions are equivalent for a right-continuous martingale $X = \{X_t, \mathcal{F}_t\}_{t\in[0,\infty)}$:

(i) It is a uniformly integrable family of random variables.

(ii) It converges in L^1, as $t \to \infty$.

(iii) It converges P-a.s. ($t \to \infty$) to an integrable random variable X_∞, such that $\{X_t, \mathcal{F}_t\}_{t\in[0,\infty]}$ is a martingale, i.e. the martingale is closed by Y.

(iv) There exists an integrable random variable Y, such that $X_t = \mathrm{E}[Y|\mathcal{F}_t]$ a.s. for all $t \geq 0$.

If (iv) holds and X_∞ is the random variable in (c), then $\mathrm{E}[Y|\mathcal{F}_\infty] = X_\infty$ a.s.

We now formulate a more general version of Doob's optional sampling theorem:

Theorem 3.22 (Optional Sampling Theorem, 2nd Version)
Let $X = \{X_t, \mathcal{F}_t\}_{t\in[0,\infty]}$ be a right-continuous submartingale (supermartingale) with a last element X_∞ and let τ_1 and τ_2 be stopping times of the filtration $\{\mathcal{F}_t\}$ with $\tau_1 \leq \tau_2$ a.s. Then

$$\mathrm{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}] \geq (\leq) X_{\tau_1}$$

Proof. See, e.g., Karatzas/Shreve (1991, pp. 19f).

Proposition 3.23 (Stopped (Sub)-Martingales)
Let $X = \{X_t, \mathcal{F}_t\}_{t\in[0,\infty)}$ be an adapted right-continuous (sub)-martingale and let τ be a stopping time. The stopped process $\{X_{t\wedge\tau}, \mathcal{F}_t\}_{t\in[0,\infty)}$ is also an adapted right-continuous (sub)-martingale.

Proof. Exercise.

Remark. The difficulty in this proposition is to show that the stopped process is a (sub)-martingale for the filtration $\{\mathcal{F}_t\}$. It is obvious that the stopped process is a (sub)-martingale for $\{\mathcal{F}_{t\wedge\tau}\}$.

Corollary 3.24 (Conditional Expectation and Stopped σ-Fields)
Let τ_1 and τ_2 be stopping times and Z be an integrable random variable. Then

$$\mathrm{E}[\mathrm{E}[Z|\mathcal{F}_{\tau_1}]|\mathcal{F}_{\tau_2}] = \mathrm{E}[\mathrm{E}[Z|\mathcal{F}_{\tau_2}]|\mathcal{F}_{\tau_1}] = \mathrm{E}[Z|\mathcal{F}_{\tau_1\wedge\tau_2}], \qquad P-a.s.$$

Proof. Exercise.

For technical reasons, we sometimes need to weaken the martingale concept.

Definition 3.16 (Local Martingale)
Let $X = \{X_t, \mathcal{F}_t\}_{t\in[0,\infty)}$ be a (continuous) process. If there exists a non-decreasing sequence $\{\tau_n\}_{n\in\mathbb{N}}$ of stopping times of $\{\mathcal{F}_t\}$, such that $\{X_t^{(n)} := X_{t\wedge\tau_n}, \mathcal{F}_t\}_{t\in[0,\infty)}$ is a martingale for each $n \geq 1$ and $P(\lim_{n\to\infty}\tau_n = \infty) = 1$, then we say that X is a (continuous) local martingale.

Remark.

- Every martingale is a local martingale (Choose $\tau_n = n$).

- However, there exist local martingales which are no martingales. Especially, $\mathrm{E}[X_t]$ need not to exist for local martingales.

Proposition 3.25
A non-negative local martingale is a supermartingale.

Proof. follows from Fatou's lemma (Exercise). □

This concept can be applied to other situations (e.g. locally bounded, locally square integrable):

Definition 3.17 (Local Property)
Let X be a stochastic process. A property π is said to hold locally if there exists a sequence of stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ with $P(\lim_{n\to\infty}\tau_n=\infty)=1$ such that $X^{(n)}$ has property π for each $n\geq 1$.

3.5 Completition, Augmentation, Usual Conditions

The following definition is very important:

Definition 3.18 (Usual Conditions)
A filtered probability space $(\Omega,\mathcal{G},P,\{\mathcal{G}_t\})$ satisfies the usual conditions if

(i) it is complete,[9]

(ii) $\mathcal{G}_0$ contains all the P-null sets of $\mathcal{G}$, and

(iii) $\{\mathcal{G}_t\}$ is right-continuous.

In the sequel, we impose the following

Standing assumption: We always consider a filtered probability space $(\Omega,\mathcal{F},P,\{\mathcal{F}_t\})$ satisfying the usual conditions.

One of the main reasons for this assumption is that the following proposition holds:

Proposition 3.26
Let $(\Omega,\mathcal{G},P,\{\mathcal{G}_t\})$ satisfy the usual conditions. Then every martingale for $\{\mathcal{G}_t\}$ possesses a unique cadlag modification.

Proof. See, e.g. Karatzas/Shreve (1991, pp. 16f).

Unfortunately, we are confronted with the following result.

Proposition 3.27
The filtration $\{\mathcal{F}^W_t\}$ is left-continuous, but fails to be right-continuous.

[9] A probability space is complete if for any $N\subset\Omega$ such that there exists an $\tilde{N}\in\mathcal{G}$ with $N\subset\tilde{N}$ and $P(\tilde{N})=0$, we have $N\in\mathcal{G}$.

Fortunately, there is a way out. Let $\mathcal{F}_\infty := \sigma\{\bigcup_{t\geq 0}\mathcal{F}_t\}$ and consider the space $(\Omega, \mathcal{F}_\infty^W, P)$. Define the *collection of P-null sets* by

$$\mathcal{N} := \{A \subset \Omega : \exists B \in \mathcal{F}_\infty^W \text{ with } A \subset B, P(B) = 0\}.$$

The P-augmentation of $\{\mathcal{F}_t^W\}$ is defined by $\mathcal{F}_t := \sigma\{\mathcal{F}_t^W \cup \mathcal{N}\}$ and is called the Brownian filtration.

Theorem 3.28 (Brownian Filtration)
The Brownian filtration, is left- and right-continuous. Besides, W is still a Brownian motion w.r.t. the Brownian filtration.

This is the reason why we work with the Brownian filtration instead of $\{\mathcal{F}_t^W\}$. For a Poisson process, a similar result holds.

Proposition 3.29 (Augmented Filtration of a Poisson Process)
The P-augmentation of the natural filtration induced by a Poisson process N is right-continuous. Besides, N is still a Poisson process w.r.t. this P-augmentation.

The proofs of the previous results can be found in Karatzas/Shreve (1991), pp. 89ff.

Remark. More generally, one can prove the following results:

(i) If a stochastic process is left-continuous, then the filtration $\{\mathcal{F}_t^X\}$ is left-continuous.

(ii) Even if X is is continuous, $\{\mathcal{F}_t^X\}$ need not to be right-continuous.

(iii) For an n-dimensional strong Markov process X the augmented filtration is right-continuous.

3.6 Ito-Integral

Our goal is to define an integral of the form:

$$\int_0^t X(s)dW(s),$$

where X is an appropriate stochastic process. From the Lebesgue-Stieltjes theory we know that for a measurable function f and a processes A of finite variation on $[0,t]$ one can define

$$\int_0^t f(s)dA(s) = \lim_{n\to\infty}\sum_{k=0}^{n-1} f\left(\tfrac{kt}{n}\right)\cdot\left(A\left(\tfrac{(k+1)t}{n}\right) - A\left(\tfrac{kt}{n}\right)\right).$$

However, the variation of a Brownian motion is unbounded on any interval. Therefore, naive stochastic integration is impossible (See Protter (2005, pp. 43ff)).

Good News: We can define a stochastic integral in terms of L^2-convergence!

Agenda. (disregarding measurability for the moment)

1st Step. Define a stochastic integral for simple processes κ (constant process except for finitely many jumps).

$$I_t(\kappa) := \int_0^t \kappa_s dW_s :=?$$

2nd Step. Every $L^2[0,T]$-process X can be approximated by a sequence $\{\kappa^{(n)}\}_n$ of simple processes such that

$$||X - \kappa^{(n)}||_T^2 := \mathrm{E}\left[\int_0^T (X(s) - \kappa^{(n)}(s))^2 \, ds\right] \to 0.$$

3rd Step. For each n the stochastic integral

$$I_t(\kappa^{(n)}) = \int_0^t \kappa^{(n)}(s) \, dW(s)$$

is a well-defined random variable and the sequence $\{I_t(\kappa^{(n)})\}_n$ converges (in some sense) to a random variable I

4th Step. Define the stochastic integral for a $L^2[0,T]$-process X as follows:

$$\int_0^t X(s) \, dW(s) := \lim_{n\to\infty} I_t(\kappa^{(n)}) = \lim_{n\to\infty} \int_0^t \kappa^{(n)}(s) \, dW(s).$$

5th Step. Extend the definition by localization.

Standing Assumption. $\{\mathcal{F}_t\}$ is the Brownian filtration of a Brownian motion W.

1st Step.

Definition 3.19 (Simple Process)
A stochastic process $\{\kappa_t\}_{t\in[0,T]}$ is a simple process if there exist real numbers $0 = t_0 < t_1 < \ldots < t_p = T$, $p \in I\!N$, and bounded random variables $\Phi_i : \Omega \to I\!R$, $i = 0, 1, \ldots, p$ such that Φ_0 is $\mathcal{F}_0$-measurable, Φ_i is $\mathcal{F}_{t_{i-1}}$-measurable, and for all $\omega \in \Omega$ we have

$$\kappa_t(\omega) = \Phi_0(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=1}^{p} \Phi_i(\omega)\mathbf{1}_{(t_{i-1},t_i]}(t)$$

Remarks.
- κ_t is $\mathcal{F}_{t_i-1}$-measurable for $t \in (t_{i-1}, t_i]$ (measurable w.r.t. the left-hand side of the interval).
- κ is a left-continuous step function.

Definition 3.20 (Stochastic Integral for Simple Processes)
For a simple process κ and $t \in [0,T]$ the stochastic integral $I(\kappa)$ is defined by

$$I_t(\kappa) := \int_0^t \kappa_s dW_s := \sum_{1\le i\le p} \Phi_i(W_{t_i\wedge t} - W_{t_{i-1}\wedge t})$$

Proposition 3.30 (Stochastic Integral for Simple Processes)
For a simple process κ we have:

(i) The integral $I(\kappa) = \{I_t(\kappa)\}_{t\in[0,T]}$ is a continuous martingale for $\{\mathcal{F}_t\}_{t\in[0,T]}$.

(ii) $\mathrm{E}\left[(I_t(\kappa))^2\right] = \mathrm{E}\left[\int_0^t \kappa_s^2 \, ds\right]$ *for all $t \in [0,T]$.*

(iii) $\mathrm{E}\left[\left(\sup_{0\le t\le T} |I_t(\kappa)|\right)^2\right] \le 4\mathrm{E}\left[\int_0^T \kappa_s^2 \, ds\right]$

Proof. (i) Exercise.

(ii) For simplicity, $t = t_{k+1}$. Then

$$\mathrm{E}[I_t(\kappa)^2] = \sum_{i=1}^{k}\sum_{j=1}^{k} \mathrm{E}\left[\Phi_i\Phi_j(W_{t_i} - W_{t_{i-1}})(W_{t_j} - W_{t_{j-1}})\right].$$

1st case: $i > j$.

$$\begin{aligned}
\mathrm{E}\left[\Phi_i\Phi_j(W_{t_i} - W_{t_{i-1}})(W_{t_j} - W_{t_{j-1}})\right] &= \mathrm{E}\left[\mathrm{E}\left[\Phi_i\Phi_j(W_{t_i} - W_{t_{i-1}})(W_{t_j} - W_{t_{j-1}})|\mathcal{F}_{t_{i-1}}\right]\right] \\
&= \mathrm{E}[\Phi_i\Phi_j(W_{t_j} - W_{t_{j-1}})(\underbrace{\mathrm{E}[W_{t_i}|\mathcal{F}_{t_{i-1}}]}_{=W_{t_{i-1}}} - W_{t_{i-1}})] \\
&= 0.
\end{aligned}$$

2nd case $i = j$.

$$\mathrm{E}\left[\Phi_i^2(W_{t_i} - W_{t_{i-1}})^2\right] = \mathrm{E}\left[\Phi_i^2\mathrm{E}[(W_{t_i} - W_{t_{i-1}})^2|\mathcal{F}_{t_{i-1}}]\right] = \mathrm{E}[\Phi_i^2(t_i - t_{i-1})]$$

Hence,

$$\mathrm{E}[I_t(\kappa)^2] = \mathrm{E}\left[\sum_{i=1}^{k}\Phi_i^2(t_i - t_{i-1})\right] = \mathrm{E}\left[\int_0^t \kappa_s^2\,ds\right].$$

(iii) Apply (i), (ii), and Doob's inequality (see Kartzas/Shreve (1991, p. 14): For a right-continuous martingale $\{M_t\}$ with $\mathrm{E}[M_t^2] < \infty$ for all $t \geq 0$ we have

$$\mathrm{E}\left[\left(\sup_{0\leq\leq T}|M_t|\right)^2\right] \leq 4\mathrm{E}[M_T^2].$$

It remains to check:

$$\mathrm{E}[I_t(\kappa)^2] = \mathrm{E}\Big[\sum_{i=1}^{k}\Phi_i^2\underbrace{(t_i - t_{i-1})}_{\leq T}\Big] \leq T\sum_{i=1}^{k}\mathrm{E}\left[\Phi_i^2\right] < \infty,$$

since, by Definition 3.19, Φ is bounded. □

Remarks.

- Stochastic integrals over $[t, T]$ are defined by
$$\int_t^T \kappa_s\,dW_s := \int_0^T \kappa_s\,dW_s - \int_0^t \kappa_s\,dW_s.$$
- Let κ_1 and κ_2 be simple processes and a, $b \in I\!R$. The stochastic integral for simple processes is linear, i.e.
$$I_t(a\kappa_1 + b\kappa_2) = aI_t(\kappa_1) + bI_t(\kappa_2).$$
This follows directly from Definition 3.20. Note that linear combinations of simple functions are simple.
- For $\kappa \equiv 1$, $\int_0^t 1\,dW_s = W_t$. Hence,
$$\mathrm{E}\left[\Big(\int_0^t dW_s\Big)^2\right] = \mathrm{E}[W_t^2] = t = \int_0^t ds.$$
This is the reason why people sometimes write $dW_t = \sqrt{dt}$, which is wrong from a mathematical point of view.

2nd Step. Every $L^2[0,T]$-process X can be approximated by a sequence $\{\kappa^{(n)}\}_n$ of simple processes.

Define the vector space

$$\begin{aligned} L^2[0,T] &:= L^2\left([0,T],\Omega,\mathcal{F},P,\{\mathcal{F}_t\}_{t\in[0,T]}\right) \\ &:= \left\{\{X_t,\mathcal{F}_t\}_{t\in[0,T]} : \text{ progressively measurable, real-valued with } \mathrm{E}\Big[\int_0^T X_s^2\,ds\Big] < \infty\right\} \end{aligned}$$

with (semi-)norm

$$||X||_T^2 := \mathrm{E}\left[\int_0^T X_s^2\,ds\right],$$

where we identify processes for which $X = Y$ only *Lebesgue* $\otimes$ P-a.s.

For simple processes the mapping $\kappa \mapsto I(\kappa)$ induces a norm on the vector space of the corresponding stochastic integrals via

$$||I(\kappa)||_{L_T}^2 := \mathrm{E}\left[\left(\int_0^T \kappa_s\,dW_s\right)^2\right] \overset{\text{Prop. 3.30 (ii)}}{=} \mathrm{E}\left[\int_0^T \kappa_s^2\,ds\right] = ||\kappa||_T^2.$$

Since the mapping $I(X)$ is linear and norm preserving, it is an isometry, called the Ito isometry.

Proposition 3.31 (Simple Processes Approximate $L^2[0,T]$-processes)
The linear subspace of simple processes is dense in $X \in L^2([0,T])$, i.e. for all $X \in L^2[0,T]$ there exists a sequence $\{\kappa^{(n)}\}_n$ of simple processes with

$$||X-\kappa^{(n)}||_T^2 = \mathrm{E}\left[\int_0^T (X_s-\kappa_s^{(n)})^2\,ds\right] \to 0 \quad (\text{ as } n\to\infty)$$

Proof. Assume that $X \in L^2[0,T]$ is continuous and bounded. Choose

$$\kappa_t^{(n)}(\omega) := X_0(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{k=0}^{2^n-1} X_{kT/2^n}(\omega)\mathbf{1}_{(kT/2^n,(k+1)T/2^n]}(t).$$

Since X is bounded, $\{\kappa^{(n)}\}_n$ is uniformly bounded and we can apply the dominated convergence theorem to obtain the desired result.

For the general case, see Karatzas/Shreve (1991, p. 133). □

3rd and 4th Step.

Define

$$\begin{aligned} \mathcal{M}_2^c([0,T]) \;\; := \;\; & \{M : M \text{ continuous martingale on } [0,T] \text{ w.r.t. } \{\mathcal{F}_t\}, \\ & \mathrm{E}[M_t^2] < \infty \text{ for every } t\in[0,T]\} \end{aligned}$$

Theorem and Definition 3.32 (Ito Integral for $X \in L^2[0,T]$)
There exists a unique linear mapping $J : L^2[0,T] \to \mathcal{M}_2^c([0,T])$ with the following properties:

(i) If κ is simple, then $P(J_t(\kappa) = I_t(X)$ for all $t \in [0,T]) = 1$.

(ii) $\mathrm{E}[J_t(X)^2] = \mathrm{E}\left[\int_0^t X_s^2\, ds\right]$ *"Ito isometry"*

J is unique in the following sense: If there exists a second mapping J' satisfying (i) and (ii) for all $X \in L^2[0,T]$, then $J(X)$ and $J'(X)$ are indistinguishable.

For $X \in L^2[0,T]$, we define

$$\int_0^t X_s\, dW_s := J_t(X)$$

for every $t \in [0,T]$ and call it the Ito integral or stochastic integral of X (w.r.t. W).

Proof. Let $X \in L^2[0,T]$.

(a) Let $\{\kappa^{(n)}\}_n$ be an approximating sequence, i.e. $||X - \kappa^{(n)}||_T \to 0$ as $n \to \infty$. Since $\{\kappa^{(n)}\}_n$ is also a Cauchy sequence w.r.t. $||\cdot||_T$, we obtain for every $t \in [0,T]$ (as n, $m \to \infty$):

$$\mathrm{E}[(I_t(\kappa^{(n)}) - I_t(\kappa^{(m)}))^2] \overset{I \text{ linear}}{=} \mathrm{E}[I_t(\kappa^{(n)} - \kappa^{(m)})^2] \overset{\text{Ito Isom.}}{=} \mathrm{E}\Big[\int_0^t (\kappa^{(n)} - \kappa^{(m)})^2 ds\Big] \to 0.$$

Hence, $\{I_t(\kappa^{(n)})\}_n$ is a Cauchy sequence in the ordinary $L^2(\Omega, \mathcal{F}_t, P)$ for random variables (t fixed!) which is a complete space and thus

$$I_t(\kappa^{(n)}) \xrightarrow{L^2} Z_t,$$

where Z_t is $\mathcal{F}_t$-measurable and $\mathrm{E}[Z_t^2] < \infty$. Since $I_t(\kappa^{(n)}) \xrightarrow{P} Z_t$ as well, there exists a subsequence $\{n_k\}_k$ such that

$$I_t(\kappa^{(n_k)}) \xrightarrow{a.s.} Z_t \tag{1}$$

as $k \to \infty$, i.e. Z_t is only P-a.s. unique (may depend on sequence!).

(b) Z **is a martingale:** Note that $I_t(\kappa^{(n)})$ is a martingale for all $n \in \mathbb{N}$. By the definition of the conditional expected value, we thus get

$$\int_A I_t(\kappa^{(n)})\, dP = \int_A I_s(\kappa^{(n)})\, dP$$

for $s < t$ and for all $A \in \mathcal{F}_s$ and $n \in \mathbb{N}$. The convergence (1) leads to (as $n \to \infty$)[10]

$$\int_A Z_t dP \longleftarrow \int_A I_t(\kappa^{(n)})\, dP = \int_A I_s(\kappa^{(n)})\, dP \longrightarrow \int_A Z_s dP.$$

Since Z is adapted, (b) is proved.

(c) By Proposition 3.26, Z possesses a right-continuous modification $\{J_t(X), \mathcal{F}_t\}_{t\in[0,T]}$ still being a square-integrable martingale. Applying Doob's inequality and the Borel-Cantelli Lemma, one can prove that $J(X)$ is even continuous (Exercise).

(d) $J_t(X)$ **is independent of the approximating sequence:** Let $\{\kappa^{(n)}\}_n$ and $\{\tilde{\kappa}^{(n)}\}_n$ be two approximating sequences for X with $I_t(\kappa^{(n)}) \xrightarrow{L^2} Z_t$ and $I_t(\tilde{\kappa}^{(n)}) \xrightarrow{L^2}$

[10] See Bauer (1992, MI, p. 92 and p. 99).

$\tilde{Z}_t$. Then $\{\hat{\kappa}^{(n)}\}_n$ with $\hat{\kappa}^{(2n)} = \kappa^{(2n)}$ (even) and $\hat{\kappa}^{(2n+1)} = \tilde{\kappa}^{2n+1}$ (odd) is an approximating sequence for X with $\{I_t(\hat{\kappa}^{(n)})\}_n$ converges in L^2 to both Z_t and $\tilde{Z}_t$. Thus they need to coincide with $J_t(X)$ P-a.s.[11]

(e) **Ito isometry:**

$$\mathrm{E}\left[J_t(X)^2\right] \overset{(*)}{=} \lim_{n\to\infty} \mathrm{E}\left[I_t(\kappa^{(n)})^2\right] = \lim_{n\to\infty} \mathrm{E}\Big[\int_0^t (\kappa_s^{(n)})^2\,ds\Big] \overset{(**)}{=} \mathrm{E}\Big[\int_0^t X_s^2\,ds\Big]$$

$(*)$ holds because $I_t(\kappa^{(n)}) \xrightarrow{L^2} J_t(X)$ (as $n\to\infty$). This is the L^2-convergence in the ordinary$L^2(\Omega,\mathcal{F}_t,P)$.

$(**)$ holds because $\kappa^{(n)} \xrightarrow{L^2} X$ (as $n\to\infty$). This is the L^2-convergence in $([0,T]\times\Omega, \mathcal{B}[0,T]\otimes\mathcal{F}, \mathit{Lebesgue}\otimes P)$.[12]

(f) **Linearity of** J follows from the linearity of I and from the fact that if $\{\kappa_X^{(n)}\}_n$ approximates X and $\{\kappa_Y^{(n)}\}_n$ approximates Y, then $\{a\kappa_X^{(n)} + b\kappa_Y^{(n)}\}_n$ approximates $aX+bY$ $(a,\, b\in I\!R)$. Therefore,

$$\begin{aligned} aJ_t(X)+bJ_t(Y) &= a\lim_{n\to\infty} I_t(\kappa_X^{(n)}) + b\lim_{n\to\infty} I_t(\kappa_Y^{(n)}) = \lim_{n\to\infty} I_t\left(a\kappa_X^{(n)} + b\kappa_Y^{(n)}\right) \\ &= J_t(aX+bY). \end{aligned}$$

(g) **Uniqueness of** J**:** Let J' be a linear mapping with properties (i) and (ii). Then

$$J'_t(X) = J'_t(\lim_{n\to\infty}\kappa^{(n)}) \overset{(*)}{=} \lim_{n\to\infty} J'_t(\kappa^{(n)}) \overset{\text{(i)}}{=} \lim_{n\to\infty} I_t(\kappa^{(n)}) = J_t(X) \quad P-a.s.$$

for all $t\in[0,T]$. Since $J(X)$ and $J'(X)$ are continuous processes, by Proposition 3.2 (ii), they are indistinguishable, i.e. $P(J_t(X) = J'_t(X)$ for all $t\in[0,T]) = 1$. The equality $(*)$ holds because linear mappings are continuous. □

5th Step. The class of admissible integrands can be extended. Define

$$\begin{aligned} H^2[0,T] &:= H^2\left([0,T],\Omega,\mathcal{F},P,\{\mathcal{F}_t\}_{t\in[0,T]}\right) \\ &:= \Big\{\{X_t,\mathcal{F}_t\}_{t\in[0,T]} : \text{ progressively measurable, real-valued with} \\ &\qquad P\Big(\int_0^T X_s^2\,ds < \infty\Big) = 1\Big\} \end{aligned}$$

In general, for $X\in H^2[0,T]$ it can happen that

$$||X||_T^2 = \mathrm{E}\left[\int_0^T X_s^2\,ds\right] = \infty.$$

Therefore, the Ito isometry needs not to hold and the approximation argument breaks down. Fortunately, one can use a localization argument:[13]

$$\begin{aligned} \tau_n(\omega) &:= T\wedge\inf\Big\{t\in[0,T] : \int_0^t X_s^2(\omega)\,ds \ge n\Big\}, \\ X_t^{(n)}(\omega) &:= X_t(\omega)\mathbf{1}_{\{\tau_n(\omega)\ge t\}}. \end{aligned}$$

[11] See Bauer (1992, MI, p. 130).

[12] Both $(*)$ and $(**)$ are a consequence of Bauer (1992, MI, p. 92).

[13] Every τ_n is a stopping time because it is the first-hitting time of $[n,\infty)$ for the continuous process $\{\int_0^t X_s^2\,ds\}_t$. See Proposition 3.11.

Hence, $X_t^{(n)}(\omega) \in L^2[0,T]$ and $J(X^{(n)})$ is defined and a martingale for fixed $n \in \mathbb{N}$. Now define

$$\boldsymbol{J_t(X) := J_t(X^{(n)})} \quad \text{for } X \in H^2[0,T] \text{ and } 0 \le t \le \tau_n.$$

Note that

- $J_t(X^{(n)}) = J_t(X^{(m)})$ for all $0 \le t \le \tau_n \wedge \tau_m$, i.e. J is well-defined,
- $\lim_{n\to\infty} \tau_n = T$ a.s. (since $P\Big(\int_0^T X_s^2\, ds < \infty\Big) = 1$, i.e. there exists an event A with $P(A) = 1$ such that for all $\omega \in A$ there exists an $c(\omega) > 0$ with $\int_0^T X_s^2(\omega)\, ds \le c(\omega)$.)

Important Results.

- By construction, for $X \in H^2[0,T]$ the process $\{J_t(X)\}_{t\in[0,T]}$ is a continuous local martingale with localizing sequence $\{\tau_n\}_n$.
- J is still linear, i.e. $J_t(aX + bY) = aJ_t(X) + bJ_t(Y)$ for a, $b \in \mathbb{R}$ and X, $Y \in H^2[0,T]$. (Choose $\tau_n := \tau_n^X \wedge \tau_n^Y$ as localizing sequence of $aX + bY$, where $\{\tau_n^X\}_n$ is the localizing sequence of X.)
- $\mathrm{E}[J_t(X)]$ needs not to exist.

Definition 3.21 (Multi-dimensional Ito Integral)
Let $\{W(t), \mathcal{F}_t\}$ be an m-dimensional Brownian motion and $\{X(t), \mathcal{F}_t\}$ an $\mathbb{R}^{n,m}$-valued progressively measurable process with all components $X_{ij} \in H^2[0,T]$. Then we define the Ito integral of X w.r.t. W as

$$\int_0^t X(s)\, dW(s) := \begin{pmatrix} \sum_{j=1}^m \int_0^t X_{1j}(s) dW_j(s) \\ \vdots \\ \sum_{j=1}^m \int_0^t X_{nj}(s) dW_j(s) \end{pmatrix}, \quad t \in [0,T],$$

where $\int_0^t X_{ij}(s) dW_j(s)$ is an one-dimensional Ito integral.

3.7 Ito's Formula

Recall our

Standing assumption: $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\})$ is a filtered probability space satisfying the usual conditions.

Additional Assumption: $\{W_t, \mathcal{F}_t\}$ is a (multi-dimensional) Brownian motion of appropriate dimension.

Definition 3.22 (Ito Process)
Let W be an m-dimensional Brownian motion.

(i) $\{X(t), \mathcal{F}_t\}$ is said to be a real-valued Ito process if for all $t \ge 0$ it possesses the representation

$$X(t) \quad = \quad X(0) + \int_0^t \alpha(s)\, ds + \int_0^t \beta(s)' dW(s)$$

$$= \quad X(0) + \int_0^t \alpha(s)\, ds + \sum_{j=1}^m \int_0^t \beta_j(s)' dW_j(s)$$

for an $\mathcal{F}_0$- measurable random variable $X(0)$, and progressively measurable processes α, β with

$$(2) \qquad \int_0^t |\alpha(s)|\, ds < \infty, \qquad \sum_{j=1}^m \int_0^t \beta_j^2(s)\, ds < \infty, \quad P - a.s. \quad \text{for all } t \geq 0.$$

(ii) An n-dimensional Ito process $X = (X^{(1)}, \ldots, X^{(n)})$ is a vector of one-dimensional Ito processes $X^{(i)}$.

Remarks.

- To shorten notations, one sometimes uses the **symbolic differential notation**

$$dX(t) = \alpha(t)dt + \beta(t)'dW(t)$$

- It can be shown that the processes α and β are unique up to indistinguishability.
- By (2), we have $\beta_j \in H^2[0, T]$.
- For a real-valued Ito process X and a real-valued progressively measurable process Y, we define

$$\int_0^t Y_s\, dX_s := \int_0^t Y_s \alpha_s\, ds + \int_0^t Y_s \beta_s'\, dW_s$$

Definition 3.23 (Quadratic Variation and Covariation)
Let $X^{(1)}$ and $X^{(2)}$ two real-valued Ito processes with

$$X^{(i)}(t) = X(0) + \int_0^t \alpha^{(i)}(s)\, ds + \int_0^t \beta^{(i)}(s)' dW(s),$$

where $i \in \{1, 2\}$. Then,

$$< X^{(1)}, X^{(2)} >_t := \sum_{j=1}^m \int_0^t \beta_j^{(1)}(s)\, \beta_j^{(2)}(s)\, ds$$

is said to be the quadratic covariation of $X^{(1)}$ and $X^{(2)}$. In particular, $< X >_t :=< X, X >_t$ is said to be the quadratic variation of an Ito process X.

Remarks. The definition may appear a bit unfounded. However, one can show the following (see Karatzas/Shreve (1991, pp. 32ff)): For fixed $t \geq 0$, let $\Pi = \{t_0, \ldots, t_n\}$ with $0 = t_0 < t_1 < \ldots < t_n = t$ be a partition of $[0, t]$. Then the p-th variation of X over the partition Π is defined by $(p > 0)$

$$V_t^{(p)}(\Pi) = \sum_{k=1}^n |X_{t_k} - X_{t_{k-1}}|^p.$$

The mesh of Π is defined by $||\Pi|| = \max_{1 \leq k \leq n} |t_k - t_{k-1}|$. For $X \in \mathcal{M}_2^c$, it holds that

$$\lim_{||\Pi|| \to 0} V_t^{(2)}(\Pi) =< X >_t \qquad \text{in probability.}$$

The main tool for working with Ito processes is Ito's formula:

Theorem 3.33 (Ito's Formula)
Let X be an n-dimensional Ito process with

$$X_i(t) = X_i(0) + \int_0^t \alpha_i(s)\,ds + \sum_{j=1}^m \int_0^t \beta_{ij}(s)\,dW_j(s).$$

Let $f : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}$ be a $C^{1,2}$-function. Then

$$\begin{aligned} f(t, X_1(t), \ldots, X_n(t)) &= f(0, X_1(0), \ldots, X_n(0)) + \int_0^t f_t(s, X_1(s), \ldots, X_n(s))\,ds \\ &\quad + \sum_{i=1}^n \int_0^t f_{x_i}(s, X_1(s), \ldots, X_n(s))\,dX_i(s) \\ &\quad + 0.5 \sum_{i,j=1}^n \int_0^t f_{x_i x_j}(s, X_1(s), \ldots, X_n(s))\,d < X_i, X_j >_s \end{aligned}$$

and all integrals are well-defined.

Proof. For simplicity, $X(0) = const.$, $n = m = 1$, and $f : \mathbb{R} \to \mathbb{R}$.

1st step. By applying a **localization** argument, we can assume that

$$M_t := \int_0^t \beta_s\,dW_s, \quad B_t := \int_0^t \alpha_s\,ds, \quad \hat{B}_t := \int_0^t |\alpha_s|\,ds,$$

$\int_0^t \beta_s^2\,ds$, and thus X are uniformly bounded by a constant C. Hence, the relevant supports of f, f', and f'' are compact implying that these functions are bounded as well. Besides, M is a martingale (since $\beta \in L^2[0,T]$).

2nd step. Let $\Pi = \{t_0, \ldots, t_p\}$ be a partition of $[0,t]$. **Taylor expension** leads to (Note: $f(X_{t_k}) = f(X_{t_{k-1}}) + f'(X_{t_{k-1}})(X_{t_k} - X_{t_{k-1}}) + \ldots$.)

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{k=1}^p (f(X_{t_k}) - f(X_{t_{k-1}})) \\ &= \underbrace{\sum_{k=1}^p f'(X_{t_{k-1}})(X_{t_k} - X_{t_{k-1}})}_{(*)} + 0.5 \underbrace{\sum_{k=1}^p f''(\eta_k)(X_{t_k} - X_{t_{k-1}})^2}_{(**)} \end{aligned}$$

with $\eta_k = X_{t_{k-1}} + \theta_k (X_{t_k} - X_{t_{k-1}})$, $\theta_k \in (0,1)$.

3rd step. Convergence of (*). We get

$$(*) = \underbrace{\sum_{k=1}^p f'(X_{t_{k-1}})(B_{t_k} - B_{t_{k-1}})}_{A_1(\Pi)} + \underbrace{\sum_{k=1}^p f'(X_{t_{k-1}})(M_{t_k} - M_{t_{k-1}})}_{A_2(\Pi)}.$$

(i) For $||\Pi|| \to 0$, we obtain (Lebesgue-Stieltjes theory)

$$A_1(\Pi) \overset{a.s., L^1}{\longrightarrow} \int_0^t f'(X_s)dB_s = \int_0^t f'(X_s)\alpha_s\,ds$$

(ii) Approximate $f'(X_s)$ by a simple process

$$Y_s^\Pi(\omega) := f'(X_0)\mathbf{1}_{\{0\}}(s) + \sum_{k=1}^p f'(X_{t_{k-1}}(\omega))\mathbf{1}_{(t_{k-1}, t_k]}(s).$$

Note that $A_2(\Pi) = \int_0^t Y_s^{\Pi}\, dM_s$.[14] By the Ito isometry,

$$\begin{aligned} \mathrm{E}\left[\int_0^t f'(X_s)\, dM_s - A_2(\Pi)\right]^2 &= \mathrm{E}\left[\int_0^t (f'(X_s) - Y_s^{\Pi})\beta_s\, dW_s\right]^2 \\ &= \mathrm{E}\left[\int_0^t (f'(X_s) - Y_s^{\Pi})^2\beta_s^2\, ds\right] \xrightarrow{L^2} 0, \end{aligned}$$

where the limit follows from the dominated convergence theorem (as $||\Pi|| \to 0$). Therefore,[15] (as $||\Pi|| \to 0$)

$$(*) = A_1(\Pi) + A_2(\Pi) \xrightarrow{L^1} \int_0^t f'(X_s)\alpha_s\, ds + \int_0^t f'(X_s)\, dM_s$$

4th step. Convergence of (**). It can be shown that (See Korn/Korn (2001, pp. 46ff))

$$(**) \xrightarrow{L^1} \int_0^t f''(X_s)\beta_s^2\, ds.$$

5th step. By steps 2-4, there exists a subsequence $\{\Pi_k\}_k$, such that the integrals converges a.s. towards the corresponding limits. Therefore, for *fixed* t both sides of Ito's formula are equal a.s. Since both sides are continuous processes in t, by Proposition 3.2 (ii), they are indistinguishable. $\square$

Remark. Often a short-hand symbolic differential notation is used ($f : [0, \infty) \times I\!R \to I\!R$)

$$df(t, X_t) = f_t(t, X_t)dt + f_x(t, X_t)dX_t + 0.5 f_{xx}(t, X_t)d < X >_t$$

Examples.

- $X_t = t$ and $f \in C^2$ leads to the fundamental theorem of integration.
- $X_t = h(t)$ and $f \in C^2$ leads to the chain rule of ordinary calculus.
- $X_t = W_t$ and $f(x) = x^2$:

$$W_t^2 = 2\int_0^t W_s\, dW_s + t$$

- $X_t = x_0 + \alpha t + \sigma W_t$, x_0, α, $\beta \in I\!R$, and $f(x) = e^x$.

$$\begin{aligned} df(X_t) &= f'(X_t)dX_t + 0.5 f''(X_t)d < X >_t \\ &= f(X_t)\Big[\alpha dt + \sigma dW_t + 0.5\sigma^2 dt\Big] \\ &= f(X_t)\Big[(\alpha + 0.5\sigma^2)dt + \sigma dW_t\Big]. \end{aligned}$$

Therefore, the solution of the stock price dynamics in the Black-Scholes model, $dS_t = S_t[\mu dt + \sigma dW_t]$, reads

$$S_t = s_0 e^{(\mu - 0.5\sigma^2)t + \sigma W_t}. \tag{3}$$

[14] This is actually a consequence of the more general definition of Ito integrals in Karatzas/Shreve (1991, pp. 129ff).

[15] L^2-convergence implies L^1-convergence

- **Ito's product rule:** Let X and Y be Ito processes and $f(x,y) = x \cdot y$ Then (omitting dependencies)

$$\begin{aligned} d(XY) &= f_x dX + f_y dY + 0.5 \underbrace{f_{xx}}_{=0} d < X > + 0.5 \underbrace{f_{yy}}_{=0} d < Y > + \underbrace{f_{xy}}_{=1} d < X, Y > \\ &= X dY + Y dX + d < X, Y > \end{aligned}$$

Definition 3.24 (Stochastic Differential Equation)
If for a given filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\})$ there exists an adapted process X with $X(0) = x \in I\!R^n$ fixed and

$$X_i(t) = x_i + \int_0^t \alpha_i(s, X(s))\, ds + \sum_{j=0}^{m} \int_0^t \beta_{ij}(s, X(s))\, dW_j(s)$$

P-a.s. for every $t \geq 0$ and $i \in \{1, \ldots, n\}$, such that

$$\int_0^t \left(|\alpha_i(s, X(s))| + \sum_{j=1}^{m} \beta_{ij}^2(s, X(s)) \right) ds < \infty$$

P-a.s. for every $t \geq 0$ and $i \in \{1, \ldots, n\}$, then X is a strong solution of the stochastic differential equation (SDE)

$$\begin{aligned} dX(t) &= \alpha(t, X(t))dt + \beta(t, X(t))dW(t) \\ X(0) &= x, \end{aligned} \tag{4}$$

where $\alpha : [0, \infty) \times I\!R^n \to I\!R^n$ and $\beta : [0, \infty) \times I\!R^n \to I\!R^{n,m}$ are given functions.

Theorem 3.34 (Existence and Uniqueness)
If the coefficients α and β are continuous functions satisfying the following global Lipschitz and linear growth conditions for fixed $K > 0$

$$\begin{aligned} ||\alpha(t,z) - \alpha(t,y)|| + ||\beta(t,z) - \beta(t,y)|| &\leq K||z - y||, \\ ||\alpha(t,y)||^2 + ||\beta(t,y)||^2 &\leq K^2(1 + ||y||^2), \end{aligned}$$

$z, y \in I\!R$, then there exists an adapted continuous process X being a strong solution of (4) and satisfying

$$\mathrm{E}[||X_t||^2] \leq C(1 + ||x||^2)e^{CT} \quad \text{for all } t \in [0, T],$$

where $C = C(K, T) > 0$. X is unique up to indistinguishability.

Proof. See Korn/Korn (2001, pp. 112ff).

In general, there do not exist explicit solutions to SDEs. However, the important family of linear SDEs admits explicit solutions:

Proposition 3.35 (Variation of Constants)
Fix $x \in I\!R$ and let A, a, B, b be progressively measurable real-valued processes with (P-a.s. for every $t \geq 0$)

$$\begin{aligned} \int_0^t (|A(s)| + |a(s)|)\, ds &< \infty, \\ \sum_{j=1}^m \int_0^t \left(B_j(s)^2 + b_j^2(s)\right)\, ds &< \infty. \end{aligned}$$

Then the linear SDE

$$\begin{aligned} dX(t) &= (A(t)X(t) + a(t))dt + \sum_{j=1}^m (B_j(t)X(t) + b_j(t))dW_j(t), \\ X(0) = x, & \end{aligned}$$

possesses an adapted Lebesgue $\otimes$ P-unique solution

$$X(t) = Z(t)\left[x + \int_0^t \frac{1}{Z(u)}\left(a(u) - \sum_{j=1}^m B_j(u)b_j(u)\right)du + \sum_{j=1}^m \int_0^t \frac{b_j(u)}{Z(u)}dW_j(u)\right],$$

where

$$(5) \qquad Z(t) = \exp\left(\int_0^t \left(A(u) - 0.5||B(u)||^2\right)du + \int_0^t B(u)'dW(u)\right).$$

is the solution of the homogeneous SDE

$$\begin{aligned} dZ(t) &= Z(t)\left[A(t)dt + B(t)'dW(t)\right], \\ Z(0) &= 1. \end{aligned}$$

Proof. For simplicity, $n = m = 1$. As in (3), one can show that Z solves the homogeneous SDE. To prove uniqueness, note that Ito's formula leads to

$$d\left(\frac{1}{Z}\right) = \frac{1}{Z}\left[(-A + B^2)dt - BdW\right].$$

Let $\tilde{Z}$ be another solution. Then, by Ito's product rule,

$$\begin{aligned} d\left(\tilde{Z}\frac{1}{Z}\right) &= \tilde{Z}d\left(\frac{1}{Z}\right) + \frac{1}{Z}d\tilde{Z} + d < \frac{1}{Z}, \tilde{Z} > \\ &= \frac{\tilde{Z}}{Z}\left[(-A + B^2)dt - BdW\right] + \frac{\tilde{Z}}{Z}\left[Adt + BdW\right] - \frac{\tilde{Z}}{Z}B^2 dt = 0 \end{aligned}$$

implying $\tilde{Z}\frac{1}{Z} = const$. Due to the initial condition, we obtain $\tilde{Z} = Z$.

To show that X satisfies the linear SDE, define $Y := X/Z$. Ito's product rule applied to YZ gives the desired result. To prove uniqueness, let X and $\tilde{X}$ be two solutions to the linear SDE. Then $\hat{X} := \tilde{X} - X$ solves the homogeneous SDE

$$d\hat{X} = \hat{X}[Adt + BdW]$$

with $\hat{X}_0 = 0$. Applying Ito's product rule to $0 \cdot Z$ shows $\hat{X}_t = 0$ a.s. for all $t \geq 0$. □

Example. In Brownian models the money market account and stocks are modeled via $(i = 1, \dots, I)$

$$
\begin{aligned}
(6) \qquad dS_0(t) &= S_0(t)r(t)dt, \quad S(0) = 1, \\
dS_i(t) &= S_i(t)\left[\mu_i(t)dt + \sum_{j=1}^{m} \sigma_{ij}(t)dW_j(t)\right], \quad S_i(0) = s_{i0} > 0.
\end{aligned}
$$

Applying “variation of constants” leads to

$$
\begin{aligned}
S_0(t) &= e^{\int_0^t r(s)\,ds}, \\
S_i(t) &= s_{i0}e^{\int_0^t \mu_i(s) - 0.5\sum_{j=1}^m \sigma_{ij}^2(s)\,ds + \sum_{j=1}^m \int_0^t \sigma_{ij}(s)dW_j(s)}.
\end{aligned}
$$

4 Continuous-time Market Model and Option Pricing

4.1 Description of the Model

We consider a continuous-time security market as given by (6).

Mathematical Assumptions.

- Recall the standing assumption: $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\})$ is a filtrered probability space satisfying the usual conditions.
- $\{W_t, \mathcal{F}_t\}$ is a (multi-dimensional) Brownian motion of appropriate dimension.
- $\{\mathcal{F}_t\}$ is the Brownian filtration.
- The coefficients r, μ, and σ are progressively measurable processes being uniformly bounded.

Notation.

- $\varphi_i(t)$: number of stocks i at time t
- $\varphi_0(t)S_0(t)$: money invested in the money market account at time t
- $X^\varphi(t) := \sum_{i=0}^I \varphi_i(t)S_i(t)$: wealth of the investor
- Assumption: $x_0 := X(0) = \sum_{i=0}^I \varphi_i(0)S_i(0) \geq 0$ (non-negative initial wealth)
- $\pi_i(t) := \varphi_i(t)S_i(t)/X(t)$: proportion invested in stock i
- $1 - \sum_i \pi_i$: proportion invested in the money market account

Definition 4.1 (Trading Strategy, Self-financing, Admissible)
(i) Let $\varphi = (\varphi_0, \dots, \varphi_I)'$ be an $I\!R^{I+1}$-valued process which is progressively measurable w.r.t. $\{\mathcal{F}_t\}$ and satisfies (P-a.s.)

$$
(7) \qquad \int_0^T |\varphi_0(t)|dt < \infty, \qquad \sum_{i=1}^I \int_0^T (\varphi_i(t)S_i(t))^2 dt < \infty.
$$

Then φ is said to be a trading strategy.

(ii) If

$$
dX(t) = \varphi(t)'dS(t)
$$

holds, then φ is said to be self-financing.

(iii) A self-financing φ is said to be admissible if $X(t) \geq 0$ for all $t \geq 0$.

Remarks.

- The corresponding portfolio process π is said to be self-financing or admissible if φ is self-financing or admissible (Recall $\pi_i = \varphi_i S_i / X$). Note that if we work with π, then we need to require that $X(t) > 0$ for all $t \geq 0$. Otherwise, π is not well-defined.
- (7) ensures that all integrals are defined.
- Compare the definition (ii) with the result of Proposition 2.11.
- Every trading strategy satisfies

$$dX(t) = d(\varphi(t)' S(t)).$$

  Hence, in abstract mathematical terms, self-financing means that φ can be put before the d operator.
- In the following, we always consider admissible trading strategies.
- Since admissible strategies are progressive, investors cannot see into the future (no clairvoyants).

Definition 4.2 (Arbitrage, Option, Replication, Fair Value)
(i) An admissible trading strategy φ is an arbitrage opportunity if (P-a.s.)

$$X^{\varphi}(0) = 0, \quad X^{\varphi}(T) \geq 0, \quad P(X^{\varphi}(T) > 0) > 0.$$

(ii) An $\mathcal{F}_T$-measurable random variable $C(T) \geq 0$ is said to be a contingent claim, $T \geq 0$.

(iii) An admissible strategy φ is a replication strategy for the claim $C(T)$ if

$$X^{\varphi}(T) = C(T) \quad P - a.s.$$

(iv) The fair time-0 value of a claim is defined as

$$C(0) = \inf\{p : D(p) \neq \emptyset\},$$

where

$$D(x) := \{\varphi : \varphi \text{ admissible and replicates } C(T) \text{ and } X^{\varphi}(0) = x\}$$

Remark. The strategy in (i) is also known as "free lottery".

Economic Assumptions.

- All investors model the dynamics of the securities as above.
- All investors agree upon σ, but may disagree upon μ.
- Investors are price takers.
- Investors use admissible strategies.

- Frictionless markets
- No arbitrage (see Definition 4.2 (i))

Proposition 4.1 (Wealth Equation)
The wealth dynamics of a self-financing strategy φ are described by the SDE

$$dX(t) = X(t)\Big[(r(t) + \pi(t)'(\mu(t) - r(t)\mathbf{1}))dt + \pi(t)'\sigma(t)dW(t)\Big], \quad X(0) = x_0,$$

the so-called wealth equation.

Proof. By definition,

$$\begin{aligned}
dX &= \varphi' dS \\
&= \varphi_0 S_0 r dt + \sum_{i=1}^{I} \varphi_i S_i \Big[\mu_i dt + \sum_{j=1}^{m} \sigma_{ij} dW_j\Big] \\
&= \Big(1 - \sum_{i=1}^{I} \pi_i\Big) X r dt + \sum_{i=1}^{I} \pi_i X \Big[\mu_i dt + \sum_{j=1}^{m} \sigma_{ij} dW_j\Big] \\
&= X\Big[\Big(r + \sum_{i=1}^{I} \pi_i(\mu_i - r)\Big)dt + \sum_{j=1}^{m}\sum_{i=1}^{I} \pi_i \sigma_{ij} dW_j\Big]
\end{aligned}$$

□

Remarks.

- Due to the boundedness of the coefficients, by "variation of constants", the condition ($P - a.s.$)
$$\int_0^t \pi_i(s)^2\, ds < \infty$$
is sufficient for the wealth equation to possess a unique solution.
- One can directly start with π (instead of φ) and require that for the wealth equation
$$\int_0^t X(s)^2 \pi_i(s)^2\, ds < \infty$$

4.2 Black-Scholes Approach: Pricing by Replication

We wish to price a European call written on a stock S with strike K, maturity T, and payoff

$$C(T) = \max\{S(T) - K; 0\},$$

where

$$dS(t) = S(t)\Big[\mu dt + \sigma dW(t)\Big]$$

with constants μ and σ. The interest rate r is constant as well.

Assumptions.

- No payouts are made to the stocks during the life-time of the call.

- There exists a $C^{1,2}$-function $f = f(t,s)$ such that the time-t price of the call is given by $C(t) = f(t, S(t))$.

Then

$$\begin{aligned} dC(t) &= f_t(t,S(t))dt + f_s(t,S(t))dS(t) + 0.5 f_{ss}(t,(S(t))d < S >_t \\ &= f_t dt + f_s S\Big(\mu dt + \sigma dW\Big) + 0.5 f_{ss} S^2\sigma^2 dt \\ &= \Big(f_t + f_s S\mu + 0.5 f_{ss} S^2\sigma^2\Big)dt + f_s S\sigma dW \end{aligned}$$

Consider a self-financing trading strategy $\varphi = (\varphi_M, \varphi_S, \varphi_C)$ and **choose** $\varphi_C(t) \equiv -1$. Then

$$\begin{aligned} dX^\varphi(t) &= \varphi_M(t)dM(t) + \varphi_S(t)dS(t) - dC(t) \\ &= \varphi_M rM dt + \varphi_S S\Big(\mu dt + \sigma dW\Big) \\ &\quad -\Big(f_t + f_s S\mu + 0.5 f_{ss} S^2\sigma^2\Big)dt - f_s S\sigma dW \\ &= \Big(\underbrace{\varphi_M rM + \varphi_S S\mu - f_t - f_s S\mu - 0.5 f_{ss} S^2\sigma^2}_{(*)}\Big)dt + \Big(\underbrace{\varphi_S S\sigma - f_s S\sigma}_{(**)}\Big)dW \end{aligned}$$

Choosing $\varphi_S(t) = f_s(t)$ leads to $(**) = 0$. Hence, the strategy is locally risk-free implying $(*) = X^\varphi(t)r$ (Otherwise there is an arbitrage opportunity with the money market account). Rewriting (*):

$$\begin{aligned} (*) &= \varphi_M rM - f_t - 0.5 f_{ss} S^2\sigma^2 \\ &= r\Big(\underbrace{\varphi_M M + \varphi_S S - C}_{=X^\varphi}\Big) \\ &\quad \underbrace{-r f_s S + rC - f_t - 0.5 f_{ss} S^2\sigma^2}_{(***)}. \end{aligned}$$

Therefore, $(***) = 0$, i.e.

$$r f_s(t,S(t))S(t) - r f(t,S(t)) + f_t(t,S(t)) + 0.5 f_{ss}(t,S(t))S^2(t)\sigma^2 = 0.$$

It follows:

Theorem 4.2 (PDE of Black-Scholes (1973))
In an arbitrage-free market the price function $f(t,s)$ of a European call satisfies the Black-Scholes PDE

$$f_t(t,s) + rs f_s(t,s) + 0.5 s^2\sigma^2 f_{ss}(t,s) - r f(t,s) = 0,$$

$(t,s) \in [0,T] \times I\!R_+$, *with terminal condition* $f(T,s) = \max\{s-K, 0\}$.

Remark. We have not used the terminal condition to derive the PDE. Therefore, every European (path-independent) option satisfies this PDE. Of course, the terminal conditions for different options are different.

Question: How should we solve this PDE?

1st approach: Merton (1973). Solve the PDE by transforming it to the heat equation $u_t = u_{xx}$ which has a well-known solution.[16]

2nd approach. Apply the Feynman-Kac theorem (see below) which states that the time-0 solution of the Black-Scholes PDE is given by

$$C(0) = \mathrm{E}\Big[\max\{Y(T) - K, 0\}\Big| Y(0) = S(0)\Big] e^{-rT},$$

where Y satisfies the SDE

$$dY(t) = Y(t)\Big[rdt + \sigma dW(t)\Big], \qquad Y(0) = S(0).$$

Theorem 4.3 (Black-Scholes Formula)
(i) The price of the European call is given by

$$C(t) = S(t) \cdot N(d_1(t)) - K \cdot N(d_2(t)) \cdot e^{-r(T-t)}$$

with

$$\begin{aligned} d_1(t) &= \frac{\ln(S(t)/K) + (r + 0.5^2)(T-t)}{\sigma\sqrt{T-t}}, \\ d_2(t) &= d_1(t) - \sigma\sqrt{T-t}. \end{aligned}$$

(ii) The trading strategy $\varphi := (\varphi_M, \varphi_S)$ with

$$\begin{aligned} \varphi_M(t) &= \frac{C(t) - f_s(t, S(t)) \cdot S(t)}{M(t)}, \\ \varphi_S(t) &= f_s(t, S(t)), \end{aligned}$$

is self-financing and a replication strategy for the call.

Proof. (i) Compute (exercise)

$$C(t) = \mathrm{E}\Big[\max\{Y(T) - K, 0\}\Big| Y(t) = S(t)\Big] e^{-r(T-t)}.$$

(ii) We have

$$X^{\varphi}(t) = \varphi_M(t) \cdot M(t) + \varphi_S(t) \cdot S(t) = C(t). \tag{8}$$

Hence, (φ_M, φ_S) replicates the call and is self-financing because

$$\begin{aligned} dX^{\varphi} &\overset{(8)}{=} dC \\ &\overset{\text{Ito}}{=} \Big(f_t + f_s S\mu + 0.5 f_{ss} S^2\sigma^2\Big)dt + f_s S\sigma dW \\ &\overset{(+)}{=} \Big(r[C - f_s S] + f_s S\mu\Big)dt + f_s S\sigma dW \\ &\overset{\text{Def. }(\varphi_M,\varphi_S)}{=} \Big(rM\varphi_M + \varphi_S S\mu\Big)dt + \varphi_S S\sigma dW \\ &= \Big(\varphi_M rM dt + \varphi_S S\Big[\mu dt + S\sigma dW\Big]\Big) \\ &= \varphi_M dM + \varphi_S dS. \end{aligned}$$

[16] See, e.g., Korn/Korn (2001, p. 109).

Note: (+) holds due to the Black-Scholes PDE

$$f_t + 0.5 f_{ss} S^2 \sigma^2 = r[C - f_s S]$$

□

Remark. Since (φ_M, φ_S) is self-financing, the strategy $(\varphi_M, \varphi_S, \varphi_C)$ with $\varphi_C = -1$ is self-financing as well because

$$d(\varphi_C(t)C(t)) = d(-C(t)) = -dC(t) = \varphi_C(t)dC(t).$$

Theorem 4.4 (Feynman-Kac Representation)
Let β and γ be functions of $(t, x) \in [0, \infty) \times I\!R$. Under technical conditions (see Korn/Korn (2001, p. 119)), there exists a unique $C^{1,2}$-solution $F = F(t, x)$ of the PDE

(9) $$F_t(t,x) + \beta(t,x)F_x(t,x) + 0.5\gamma^2(t,x)F_{xx}(t,x) - rF(t,x) = 0$$

with terminal condition $F(T, x) = h(x)$ possessing the representation

$$F(t_0, x) = e^{-r(T-t_0)}\mathrm{E}[h(X(T))|X(t_0) = x],$$

where X satisfies the SDE

$$dX(s) = \beta(s, X(s))ds + \gamma(s, X(s))dW(s)$$

with initial condition $X(t_0) = x$.

Sketch of Proof. (i) Firstly assume $r = 0$. If F is a $C^{1,2}$-function, we can apply Ito's formula

$$dF = F_t dt + F_x dX + 0.5 F_{xx} d < X > .$$

Since $d < X >= \gamma^2 dt$, we get

$$\begin{aligned} dF &= F_t dt + F_x\Big(\beta dt + \gamma dW\Big) + 0.5\gamma^2 F_{xx} dt \\ &= \Big(F_t + \beta F_x + 0.5\gamma^2 F_{xx}\Big)dt + \gamma F_x dW. \end{aligned}$$

Suppose that there exists a function F satisfying the PDE, i.e.

$$F_t + \beta F_x + 0.5\gamma^2 F_{xx} = 0$$

with $F(T, x) = h(x)$. Then

$$dF = \gamma F_x dW$$

or in integral notation

$$F(s, X(s)) = F(t_0, X(t_0)) + \int_{t_0}^{s} \gamma(u, X(u))F_x(u, X(u))dW(u).$$

If $\mathrm{E}\big[\int_{t_0}^{s} \gamma(u, X_u)^2 F_x(u, X_u)^2\, du\big] < \infty$, the stochastic integral has zero expectation and thus

$$F(t_0, x) = \mathrm{E}[F(s, X(s)|X(t_0) = x].$$

Choose $s = T$ and note that $F(T, X(T)) = h(X(T))$.

(ii) For $r \neq 0$, multiply the PDE (9) by e^{-rt}:

$$F_t(t,x)e^{-rt} + \beta(t,x)F_x(t,x)e^{-rt} + 0.5\gamma^2(t,x)F_{xx}(t,x)e^{-rt} - rF(t,x)e^{-rt} = 0.$$

Set $G(t,x) := F(t,x)e^{-rt}$. Since

$$G_t(t,x) = F_t(t,x)e^{-rt} - rF(t,x)e^{-rt},$$

we get

$$G_t(t,x) + \beta(t,x)G_x(t,x) + 0.5\gamma^2(t,x)G_{xx}(t,x) = 0$$

with terminal condition $G(T,x) = e^{-rT}h(x)$. Now we can apply (i):

$$G(t_0,x) = \mathrm{E}[e^{-rT}h(X(T))|X(t_0) = x].$$

Substituting $G(t_0,x) = e^{-rt_0}F(t_0,x)$ gives the desired result. □

Remarks.

- Two crucial points are not proved: Firstly, we have not checked under which conditions a $C^{1,2}$-function F exists. Secondly, we have not checked under which conditions $\gamma(\cdot,X)F_x(\cdot,X) \in L^2[0,T]$.
- To solve the Black-Scholes PDE, choose $\beta(t,x) = rx$ and $\gamma(t,x) = x\sigma$.
- The Feynman-Kac representation can be generalized to a multidimensional setting, i.e. X is a multidimensional Ito process, with stochastic short rate $r = r(X)$.

4.3 Pricing with Deflators

In a single-period model the deflator is given by (See remark to Proposition 2.7)

$$H = \frac{1}{1+r}\frac{dQ}{dP}$$

This motivates the definition of the deflator in continuous-time:

$$H(t) := \frac{1}{M(t)}Z(t),$$

where

$$\begin{aligned} M(t) &:= \exp\left(\int_0^t r(s)\,ds\right), \\ Z(t) &:= \exp\left(-0.5\int_0^t ||\theta(s)||^2\,ds - \int_0^t \theta(s)'\,dW(s)\right), \end{aligned}$$

The process θ is implicitly defined as the solution to the following system of linear equations:

(10) $$\sigma(t)\theta(t) = \mu(t) - r(t)\mathbf{1}.$$

Applying Ito's product rule yields the following SDE for H:

$$dH(t) = d(Z(t)M(t)^{-1}) = -H(t)[r(t)dt + \theta(t)'dW(t)]$$

Remarks.

- In this section we are going to work under the **physical measure**. Nevertheless we will see later on that Z can be interpreted as a density inducing a change of measure (Girsanov's Theorem).
- The system of equations (10) needs not to have a (unique) solution.
- The process θ is said to be the market price of risk and can be interpreted as a risk premium (excess return per units of volatility).

The following theorem shows that H has indeed the desired properties:

Theorem 4.5 (Arbitrage-free Pricing with Deflator and Completeness)
(i) Assume that (10) has a bounded solution for θ. Let φ be an admissible portfolio strategy. Then $\{H(t)X^{\varphi}(t)\}_{t\in[0,T]}$ is both a local martingale and a supermartingale implying

$$\mathrm{E}[H(t)X^{\varphi}(t)] \le X^{\varphi}(0) \qquad \text{for all } t \in [0,T]. \tag{11}$$

Besides, the market is arbitrage-free.

(ii) Assume that $I = m$ ("# stocks = # sources of risk") and that σ is uniformly positive definite.[17] Then (10) has a unique solution which is uniformly bounded. Furthermore, let $C(T)$ be an $\mathcal{F}_T$-measurable contingent claim with

$$x := \mathrm{E}[H(T)C(T)] < \infty \tag{12}$$

Then there exists a $P \otimes$ Lebesgue-unique portfolio process φ such that φ is a replication strategy for $C(T)$, i.e. (P-a.s.)

$$X^{\varphi}(T) = C(T),$$

and $X^{\varphi}(0) = x$. Hence, the market is complete. The time-t price of the claim equals

$$C(t) = \frac{\mathrm{E}[H(T)C(T)|\mathcal{F}_t]}{H(t)},$$

i.e. the process $\{H(t)C(t)\}_{t\in[0,T]}$ is a martingale.

Remarks.
- The boundedness assumption in (i) can be weakened.
- The price in (ii) is thus equal to the expected discounted payoff under the physical measure where we discount with the state-dependent discount factor H, the deflator.
- Compare (ii) with Theorem 2.13.

Proof. For simplicity, $I = 1$. (i) Applying Ito's product rule, one gets for an admissible $\varphi = (\varphi_0, \varphi_1)$:[18]

$$d(HX^{\varphi}) \quad = \quad HdX^{\varphi} + X^{\varphi}dH + d < H, X^{\varphi} >$$

[17] The process σ is uniformly positive definite if there exists a constant $K > 0$ such that $x'\sigma(t)\sigma(t)'x \ge Kx'x$ for all $x \in I\!R^m$ and all $t \in [0,T]$, P-a.s. This requirement implies that $\sigma(t)$ is regular.

[18] Note $\theta\sigma = \mu - r$.

$$\begin{aligned} &= HX^{\varphi}rdt + H\varphi_1 S(\mu - r)dt + H\varphi_1 S\sigma dW - X^{\varphi}Hrdt \\ &\quad -X^{\varphi}H\theta dW - H\varphi_1 S\theta\sigma dt \\ &= H(\varphi_1 S\sigma - X^{\varphi}\theta)dW. \end{aligned} \tag{13}$$

Since φ is admissible, we have $X^{\varphi}(t) \geq 0$. By definition, $H(t) \geq 0$ and thus $H(t)X(t) \geq 0$. By Proposition 3.25, HX is a supermartingale which proves (11). Note that an arbitrage strategy would violate this supermartingale property.

(ii) Clearly, $\theta = (\mu - r)/\sigma$ is unique and uniformly bounded by some $K_\theta > 0$.[19] Define

$$X(t) := \frac{\mathrm{E}[H(T)C(T)|\mathcal{F}_t]}{H(t)}$$

By definition, $X(t)$ is adapted, $X(0) = x$, $X(t) \geq 0$, and $X(T) = C(T)$ (P-a.s.).[20] Set $M(t) := H(t)X(t) = \mathrm{E}[H(T)C(T)|\mathcal{F}_t]$. Due to (12), M is a martingale with $M(0) = x$. Applying the martingale representation theorem, there exists an $\{\mathcal{F}_t\}$-progressively measurable real-valued process ψ with $\int_0^T \psi(s)^2\,ds < \infty$ such that $(P - a.s.)$

$$H(t)X(t) = M(t) = x + \int_0^t \psi(s)\,dW(s). \tag{14}$$

Since H and $\int_0^{\cdot} \psi(s)\,dW(s)$ are continuous, we can choose X to be continuous. Comparing (13) and (14) yields[21]

$$H(\varphi_1 S\sigma - X\theta) = \psi.$$

Since there exists a $K_\sigma > 0$ such that $\sigma(t) \geq K_\sigma$ for all $t \in [0, T]$,[22] we get

$$\varphi_1 = \frac{\psi/H + X\theta}{S\sigma}.$$

Since $X^{\varphi} = \varphi_0 M + \varphi_1 S$ for any strategy φ, we define

$$\varphi_0 := \frac{X - \varphi_1 S}{M}$$

which implies that φ is self-financing and $X = X^{\varphi}$. Since $M > 0$, $H > 0$, and $X \geq 0$ are continuous, M as well as H are pathwise bounded away from zero and X is pathwise bounded from above, i.e. $M(\omega) \geq K_M(\omega) > 0$, $H(\omega) \geq K_H(\omega) > 0$, and $X(\omega) \leq K_X(\omega) < \infty$ for a.a. $\omega \in \Omega$. Therefore, the process is a trading strategy because[23]

$$\begin{aligned} \int_0^T (\varphi_1(s)S(s))^2\,ds &= \int_0^T \left(\frac{\psi(s)/H(s) + X(s)\theta(s)}{\sigma(s)}\right)^2 ds \\ &\leq \frac{2}{K_\sigma^2}\left(\frac{1}{K_H^2}\int_0^T \psi(s)^2\,ds + K_X^2 K_\theta^2 T\right) < \infty, \end{aligned}$$

[19] Note that all coefficients are bounded and σ is uniformly positive definite.

[20] Since $\mathcal{F}_0$ is the trivial σ-algebra augemented by null sets, is every $\mathcal{F}_0$-measurable random variable a.s. constant and the conditioned expected value is equal to the unconditioned expected value.

[21] Note that the representation of an Ito process is unique.

[22] By assumption, σ is uniformly positive definite.

[23] $(a + b)^2 \leq 2a^2 + 2b^2$.

$$\begin{aligned}\int_0^T |\varphi_0(s)|\,ds &= \int_0^T \frac{|X(s)-\varphi_1(s)S(s)|}{M(s)}\,ds \\ &\leq \frac{1}{K_M}\left(K_X T + \int_0^T 1 + (\varphi_1(s)S(s))^2\,ds\right) < \infty\end{aligned}$$

for P-a.s. □

Theorem 4.6 (Martingale Representation Theorem)
Let $\{M_t\}_{t\in[0,T]}$ be a real-valued process being adapted to the Brownian filtration.

(i) If M is a square integrable martingale, i.e. $\mathrm{E}[M_t^2] < \infty$ for all $t \in [0,T]$, then there exists a progressively measurable $I\!R^m$-valued process $\{\psi_t\}_{t\in[0,T]}$ with

$$\mathrm{E}\left[\int_0^T ||\psi_s||^2\,ds\right] < \infty \tag{15}$$

and

$$M_t = M_0 + \int_0^t \psi_s' dW_s, \quad P\text{-a.s.}$$

(ii) If M is a local martingale, then the result from (i) holds with

$$\int_0^T ||\psi_s||^2\,ds < \infty$$

instead of (15).

In both cases, ψ is $P \otimes$ Lebesgue-unique.

Proof. See Korn/Korn (2001, pp. 71ff).

Remarks.
- All Brownian (local) martingales are stochastic integrals and vice versa!
- The quadratic (co)variation is defined for all (local) Brownian martingales.

Proposition and Definition 4.7 (Quadratic (Co)variation)
(i) Let X, Y be Brownian local martingales. The quadratic covariation $< X, Y >$ is the unique adapted, continuous process of bounded variation with $< X, Y >_0 = 0$ such that $XY- < X, Y >$ is a Brownian local martingale. If $X = Y$, this process is non-decreasing.

(ii) Let $(\Omega, \mathcal{G}, P, \{\mathcal{G}_t\})$ be a filtered probability space satisfying the usual conditions and let X, Y be continuous local martingales for the filtration $\{\mathcal{G}_t\}$. Then the result of (i) holds with $XY- < X, Y >$ being a local martingale for $\{\mathcal{G}_t\}$.

Proof. (i) By the martingale representation theorem, $dX = \alpha dW$ and $dY = \beta dW$ with progressively measurable processes α and β. Ito's product rule yields

$$d(XY) = YdX + XdY + d < X, Y >= Y\alpha dW + X\beta dW + \alpha\beta dt.$$

Since Ito integrals are local Brownian martingales, $XY- < X, Y >$ is a Brownian martingale. The rest is obvious.

(ii) See Kartzas/Shreve (1991, p. 36). □

Remark. One can show that if X and Y are square-integrable, then $XY - < X, Y >$ is a martingale.

4.4 Pricing with Risk-neutral Measures

The deflator is defined as

$$H(t) = \frac{1}{M(t)} Z(t),$$

where

$$dZ(t) = -Z(t)\theta(t)'\, dW(t)$$

is a local martingale. If it is even a martingale, we have $\mathrm{E}[Z(T)] = 1$ and we can define a probability measure on $\mathcal{F}_T$ by

$$Q(A) := \mathrm{E}[\mathbf{1}_A Z(T)], \quad A \in \mathcal{F}_T.$$

Hence, $Z(T)$ is the density of Q and we can rewrite the result from Theorem 4.5 as

$$C(0) = \mathrm{E}[H(T)C(T)] = \mathrm{E}\left[Z(T)\frac{C(T)}{M(T)}\right] = \mathrm{E}_Q\left[\frac{C(T)}{M(T)}\right],$$

i.e. the price of a claim equals its expected discounted payoff where we discount with the money market account (instead of the deflator) and expectation is calculated w.r.t. the measure Q (instead of the physical measure).

Two important **questions** arise:

Q1. Under which conditions is Z a martingale? (Novikov condition)

Q2. What happens to the dynamics of the assets if we change to the measure Q? (Girsanov's theorem)

Proposition 4.8
Let θ be progressive and $K > 0$. If $\int_0^T ||\theta(s)||^2\, ds \leq K$, then the process $\{Z(t)\}_{t\in[0,T]}$ defined by

$$dZ(t) = -Z(t)\theta(t)'dW(t), \quad Z(0) = 1, \tag{16}$$

is a martingale.

Proof. For simplicity, $m = 1$. The process Z is a local martingale. By variations of constants, we get

$$Z(t) = \exp\left(-0.5\int_0^t \theta(s)^2\, ds - \int_0^t \theta(s)\, dW(s)\right) \geq 0. \tag{17}$$

By Proposition 3.25, it is thus a supermartingale. Hence, it is sufficient to show that $\mathrm{E}[Z(T)] = 1$. Define

$$\tau_n := \inf\Big\{t \geq 0 : \Big|\int_0^t \theta(s)\, dW(s)\Big| \geq n\Big\}.$$

Due to (17), for fixed $n \in I\!N$, we get $Z(T \wedge \tau_n)^2 \leq e^{2n}$ and thus $\int_0^{T\wedge\tau_n} Z(s)^2\theta(s)^2\,ds \leq K\,e^{2n}$. From (16), we obtain $\mathrm{E}[Z(T \wedge \tau_n)] = 1$. Therefore, the claim is proved if $\{Z(T \wedge \tau_n)\}_{n \in I\!N}$ is UI because in this case

$$\lim_{n\to\infty} \mathrm{E}[Z(T \wedge \tau_n)] = \mathrm{E}[Z(T)].$$

We obtain the estimate

$$|Z(T \wedge \tau_n)|^{1+\varepsilon} \leq \underbrace{\exp\Big(0.5(\varepsilon+\varepsilon^2)\int_0^T \theta(s)^2\,ds\Big)}_{\text{independent of } n \text{ and bounded by assumption}} Y(T \wedge \tau_n),$$

where

$$dY(t) = -Y(t)(1+\varepsilon)\theta(s)dW(s), \quad Y(0) = 1.$$

With the same arguments leading to $\mathrm{E}[Z(T\wedge\tau_n)] = 1$ we conclude $\mathrm{E}[Y(T\wedge\tau_n)] = 1$. Therefore,

$$\sup_{n \in I\!N} \mathrm{E}|Z(T \wedge \tau_n)|^{1+\varepsilon} < \infty,$$

which, by Proposition 3.19, proves UI of $\{Z(T \wedge \tau_n)\}_{n\in I\!N}$. □

Remark. The requirement in the proposition is a strong one. It can be replaced by the Novikov condition (See Karatzas/Shreve (1991, pp. 198ff)):

$$\mathrm{E}\Big[\exp\Big(0.5\int_0^T ||\theta(s)||^2\,ds\Big)\Big] < \infty.$$

We turn to the second question and define ($Q := Q_T$)

$$Q_t(A) := \mathrm{E}[\mathbf{1}_A Z(t)], \quad A \in \mathcal{F}_t, \quad t \in [0,T]. \tag{18}$$

The following lemma shows that Q_t is well-defined.

Lemma 4.9
For a bounded stopping time $\tau \in [0,T]$ and $A \in \mathcal{F}_\tau$, we get $Q_T(A) = Q_\tau(A)$.

Proof. The optional sampling theorem (OS) yields (See Theorem 3.17)

$$Q_T(A) = \mathrm{E}[\mathbf{1}_A Z(T)] = \mathrm{E}[\mathrm{E}[\mathbf{1}_A Z(T)|\mathcal{F}_\tau]] = \mathrm{E}[\mathbf{1}_A \mathrm{E}[Z(T)|\mathcal{F}_\tau]] \overset{OS}{=} \mathrm{E}[\mathbf{1}_A Z(\tau)] = Q_\tau(A).$$

□

To prove Girsanov's theorem, we need some additional tools.

Theorem 4.10 (Levy's Characterization of the Brownian Motion)
An m-dimensional stochastic process X with $X(0) = 0$ is an m-dimensional Brownian motion if and only if it is a continuous local martingale with[24]

$$< X_j, X_k >_t = \delta_{jk} t.$$

[24] Note that $\delta_{ij} = \mathbf{1}_{\{i=j\}}$ is the Kronecker delta.

Proof. For simplicity $m = 1$. Clearly, the Brownian motion is a continuous local martingale with $< W >_t = t$. To prove the converse, fix $u \in I\!R$ and set $f(t,x) = \exp(iux + 0.5u^2t)$ and $Z_t := f(t, X_t)$, where $i = \sqrt{-1}$. Since $f \in C^{1,2}$ is analytic, a complex-valued version of Ito's formula yields

$$\begin{aligned} dZ &= f_t dt + f_x dX + 0.5 f_{xx} d < X > \\ &= 0.5u^2 f dt + iuf dX - 0.5u^2 f d < X > \\ &\overset{<X>_t = t}{=} iuf dX. \end{aligned}$$

Since an integral w.r.t. a local martingale is a local martingale as well,[25] Z is a (complex-valued) local martingale which is bounded on $[0,T]$ for every $T \in I\!R$. Hence, Z is a martingale implying $\mathrm{E}[\exp(iuX_t + 0.5u^2t|\mathcal{F}_s] = \exp(iuX_s + 0.5u^2 s)$, i.e.

$$\mathrm{E}[\exp(iu(X_t - X_s))|\mathcal{F}_s] = \exp(-0.5u^2(t-s)).$$

Since this hold for any $u \in I\!R$, the difference $X_t - X_s \sim \mathcal{N}(0, t-s)$ and $X_t - X_s$ is independent of $\mathcal{F}_s$. □

Proposition 4.11 (Bayes Rule)
Let μ and ν be two probability measures on a measurable space $(\Omega, \mathcal{G})$ such that $d\nu/d\mu = f$ for some $f \in L^1(\mu)$. Let X be a random variable on $(\Omega, \mathcal{G})$ such that

$$\mathrm{E}_\nu[|X|] = \int |X| f d\mu < \infty.$$

Let $\mathcal{H}$ be a σ-field with $\mathcal{H} \subset \mathcal{G}$. Then

$$\mathrm{E}_\nu[X|\mathcal{H}]\,\mathrm{E}_\mu[f|\mathcal{H}] = \mathrm{E}_\mu[fX|\mathcal{H}] \quad a.s.$$

Proof. Let $H \in \mathcal{H}$. We start with the left-hand side:

$$\begin{aligned} \int_H \mathrm{E}_\mu[fX|\mathcal{H}]\, d\mu &= \int_H X f\, d\mu = \int_H X d\nu = \int_H \mathrm{E}_\nu[X|\mathcal{H}]\, d\nu = \int_H \mathrm{E}_\nu[X|\mathcal{H}] f\, d\mu \\ &= \mathrm{E}_\mu\big[\mathrm{E}_\nu[X|\mathcal{H}] \cdot f \mathbf{1}_H\big] = \mathrm{E}_\mu[\mathrm{E}_\mu[\mathrm{E}_\nu[X|\mathcal{H}] \cdot f \mathbf{1}_H|\mathcal{H}]] \\ &= \mathrm{E}_\mu[\mathbf{1}_H \mathrm{E}_\nu[X|\mathcal{H}] \cdot \mathrm{E}_\mu[f|\mathcal{H}]] = \int_H \mathrm{E}_\nu[X|\mathcal{H}] \cdot \mathrm{E}_\mu[f|\mathcal{H}]\, d\mu. \end{aligned}$$

□

Lemma 4.12
Let $\{Y_t, \mathcal{F}_t\}_{t \in [0,T]}$ be a right-continuous adapted process. The following statements are equivalent:

(i) $\{Z_t Y_t, \mathcal{F}_t\}_{t \in [0,T]}$ is a local P-martingale.

(ii) $\{Y_t, \mathcal{F}_t\}_{t \in [0,T]}$ is a local Q_T-martingale.

[25] For instants, a Brownian local martingale can be represented as Ito integral.

Proof. Assume (i) to hold and let $\{\tau_n\}_n$ be a localizing sequence of ZY. Then[26]

$$\begin{aligned} \mathrm{E}_{Q_T}[Y_{t\wedge\tau_n}|\mathcal{F}_s] &\overset{\text{Lem 4.9}}{=} \mathrm{E}_{Q_{t\wedge\tau_n}}[Y_{t\wedge\tau_n}|\mathcal{F}_s] \overset{\text{Bayes}}{=} \frac{\mathrm{E}_P[Z_{t\wedge\tau_n}Y_{t\wedge\tau_n}|\mathcal{F}_s]}{\mathrm{E}_P[Z_{t\wedge\tau_n}|\mathcal{F}_s]} \\ &= \frac{Z_{s\wedge\tau_n}Y_{s\wedge\tau_n}}{Z_{s\wedge\tau_n}} = Y_{s\wedge\tau_n}. \end{aligned}$$

The converse implication can be proved similarly (exercise). □

Theorem 4.13 (Girsanov's Theorem)
We assume that Z is a martingale and define the process $\{W^Q(t),\mathcal{F}_t\}$ by

$$W_j^Q(t) := W_j(t) + \int_0^t \theta_j(s)\,ds, \quad t \geq 0,$$

$j \in \{1,\ldots,m\}$. For fixed $T \in [0,\infty)$, the process $\{W^Q(t),\mathcal{F}_t\}_{t\in[0,T]}$ is an m-dimensional Brownian motion on $(\Omega,\mathcal{F}_T,Q_T)$, where Q_T is defined by (18).

Proof. For simplicity, $m = 1$. By Levy's characterization, we need to show that

(i) $\{W^Q(t),\mathcal{F}_t\}_{t\in[0,T]}$ is a continuous local martingale on $(\Omega,\mathcal{F}_T,Q_T)$.

(ii) $< W^Q >_t^Q = t$ on $(\Omega,\mathcal{F}_T,Q_T)$.

ad (i). Ito's product rule yields

$$\begin{aligned} d(ZW^Q) &= ZdW^Q + W^Q dZ + d < W^Q, Z > \\ &= ZdW + Z\theta dt - W^Q Z\theta dW - Z\theta dt \\ &= (Z - W^Q Z\theta)dW, \end{aligned}$$

i.e. ZW^Q is a local P-martingale. By Lemma 4.12, the process W^Q is a local Q_T-martingale as well.

ad (ii). Ito's product rule leads to

$$\begin{aligned} d\Big(Z\cdot((W^Q)^2 - t)\Big) &= d\Big(W^Q\cdot(ZW^Q)\Big) - d(Z\cdot t) \\ &= W^Q d(ZW^Q) + ZW^Q dW^Q + d < ZW^Q, W^Q > - Zdt - tdZ \\ &= W^Q d(ZW^Q) + ZW^Q(\theta dt + dW) + (Z - W^Q Z\theta)dt - Zdt + t\theta Z dW \\ &= W^Q d(ZW^Q) + Z(W^Q + t\theta)dW, \end{aligned}$$

i.e. $\{Z_t\cdot((W_t^Q)^2 - t)\}$ is a local P-martingale. Applying Lemma 4.12 again yields that $\{(W_t^Q)^2 - t\}$ is a local Q_T-martingale. By Proposition 4.7 (ii), the claim (ii) follows. □

Corollary 4.14 (Girsanov II)
If $\{M_t\}_{t\in[0,T]}$ is a Brownian local P-martingale and

$$D_t := \int_0^t \frac{1}{Z_s} d < Z, M >_s,$$

[26] Note that $Y_{t\wedge\tau_n}$ is $\mathcal{F}_{t\wedge\tau_n}$-measurable, i.e. the conditional expected value can be computed w.r.t. $Q_{t\wedge\tau_n}$ instead of Q_T. Lemma 4.9 ensures that both measures coincide on $\mathcal{F}_{t\wedge\tau_n}$. Besides, $\mathcal{F}_{t\wedge\tau_n} \subset \mathcal{F}_T$, i.e. Bayes rule can be applied. Furthermore, by Proposition 3.23, $\mathrm{E}_P[Z_{t\wedge\tau_n}|\mathcal{F}_s] = Z_{s\wedge\tau_n}$.

then $\{M_t - D_t\}_{t\in[0,T]}$ is a continuous local Q_T-martingale and $< M - D >^Q = < M >$, i.e. the quadratic variations of $M - D$ under Q_T and M under P coincide.

Proof. By the martingale representation theorem, we have $dM = \alpha dW$ for some progressive α. Hence,

$$dD = \frac{1}{Z} d < Z, M > = -\frac{1}{Z}\theta Z \alpha dt = -\theta\alpha dt$$

and thus

$$d(M - D) = \alpha dW + \theta\alpha dt = \alpha dW^Q.$$

Therefore, $M - D$ is a local Q-martingale. Besides,

$$d < M - D >_t^Q = \alpha^2 dt = d < M >_t,$$

which completes the proof. □

Theorem 4.15 (Converse Girsanov)
Let Q be an equivalent probability measure to P on $\mathcal{F}_T$. Then the density process $\{Z(t), \mathcal{F}_t\}_{t\in[0,T]}$ defined by

$$Z(t) := \left.\frac{dQ}{dP}\right|_{\mathcal{F}_t}$$

is a positive Brownian P-martingale possessing the representation

$$dZ(t) = -Z(t)\theta(t)dW(t), \quad Z(0) = 1,$$

for a progressively measurable m-dimensional process θ with

$$\int_0^T ||\theta(s)||^2 \, ds < \infty \quad P - a.s. \tag{19}$$

Proof. Since Q and P are equivalent on $\mathcal{F}_T$, both measures are equivalent on $\mathcal{F}_t$, $t \in [0, T]$, i.e. $Z(t) > 0$ for all $t \in [0, T]$. Besides, for $A \in \mathcal{F}_t$

$$\int_A Z(T)dP = \int_A dQ/dP|_{\mathcal{F}_T} \overset{\text{Lem 4.9}}{=} \int_A dQ/dP|_{\mathcal{F}_t} = \int_A Z(t)dP,$$

i.e. $Z(t) = \mathrm{E}[Z(T)|\mathcal{F}_t]$ implying that Z is a Brownian martingale. By the martingale representation theorem,

$$Z(t) = 1 + \int_0^t \psi(s)' dW(s)$$

with a progressive process ψ satisfying $\int_0^T ||\psi(s)||^2 \, ds < \infty$, $P - a.s.$. Since $Z > 0$, we have $Z(\omega) \geq K_Z(\omega) > 0$ for a.a. $\omega \in \Omega$. Defining

$$\theta(t) = -\frac{\psi(t)}{Z(t)}$$

thus implies (19) which completes the proof. □

Definition 4.3 (Doleans-Dade Exponential)
For an Ito process X with $X_0 = 0$ the Doleans-Dade exponential (stochastic exponential) of X, written $\mathcal{E}(X)$, is the unique Ito process Z that is the solution of

$$Y_t = 1 + \int_0^t Y_s dX_s.$$

Remarks.

- Therefore, the density Z is sometimes written as $\mathcal{E}(\int_0^\cdot \theta_s dW_s)$.
- One can show that $\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + < X, Y >)$ and that $(\mathcal{E}(X))^{-1} = \mathcal{E}(-X + < X, X >)$.

We now come back to our pricing problem. We set $Q := Q_T$. Recall our definitions

$$H(t) := \frac{1}{M(t)} Z(t),$$

where

$$\begin{aligned} M(t) &:= \exp\left(\int_0^t r(s)\, ds\right), \\ Z(t) &:= \exp\left(-0.5 \int_0^t ||\theta(s)||^2\, ds - \int_0^t \theta(s)'\, dW(s)\right), \end{aligned}$$

and θ is implicitly defined by:

$$\sigma(t)\theta(t) = \mu(t) - r(t)\mathbf{1}. \tag{20}$$

For bounded θ, by Girsanov's theorem, $Z(t) = dQ/dP|_{\mathcal{F}_t}$ is a density process for the measure $Q_t(A) = \mathrm{E}[Z(t)\mathbf{1}_A]$ and $\{W^Q(t), \mathcal{F}_t\}$ defined by

$$W_j^Q(t) := W_j(t) + \int_0^t \theta_j(s)\, ds, \quad t \geq 0,$$

is a Q-Brownian motion.

Definition 4.4 (Equivalent Martingale Measure)
A probability measure Q is a (weak) equivalent martingale measure if

(i) $Q \sim P$ and

(ii) all discounted price processes S_i/M, $i = 1, \ldots, I$, are (local) Q-martingales.

Remark. An equivalent martingale measure (EMM) is also said to be a risk-neutral measure.

Proposition 4.16
If Q is a weak EMM and

$$\mathrm{E}\left[\exp\left(0.5 \int_0^T ||\sigma_i(s)||^2\, ds\right)\right] < \infty \tag{21}$$

for every $i \in \{1, \ldots, I\}$, then Q is already an EMM.

Proof. Define $Y_i := S_i/M$. Since Q is a weak martingale measure, Y_i is a local Q-martingale. By the martingale representation theorem and the uniqueness of the representation of an Ito process,

$$dY_i = Y_i\sigma_i' dW^Q.$$

By the Novikov condition (21), this is a Q-martingale. □

Remark. Since we have assumed that σ is uniformly bounded, we need not to distinguish between EMM and weak EMM.

Theorem 4.17 (Characterization of EMM)
In the Brownian market (6), the following are equivalent:

(i) The probability measure Q is an EMM.

(ii) There exists a solution θ to the system of equations (20) such that the process $Z = \mathcal{E}(\int_0^\cdot \theta_s dW_s)$ is a Girsanov density inducing the probability measure Q.

Proof. Define $Y_i := S_i/M$. Assume (i) to hold. Since Q is a martingale measure,

$$dY_i = Y_i\sigma_i' dW^Q. \tag{22}$$

Besides, since Q is equivalent to P, by the converse Girsanov, there exists a process θ such that $dQ/dP = \mathcal{E}(\int_0^\cdot \theta_s' dW_s)$ and $dW^Q = dW + \theta dt$ is a Brownian increment under Q. Therefore,

$$dS_i = S_i[\mu_i dt + \sigma_i' dW] = S_i[(\mu_i - \sigma_i'\theta)dt + \sigma_i' dW^Q]$$

implying

$$dY_i = Y_i[(\mu_i - \sigma_i'\theta - r)dt + \sigma_i' dW^Q] \tag{23}$$

Comparing (22) and (23) gives (ii).

On the other hand, assume (ii) to hold. By definition, Q is equivalent to P and $dW^Q = dW + \theta dt$ is a Q-Brownian increment. Ito's formula leads to

$$dY_i = Y_i[(\mu_i - r)dt + \sigma_i' dW] = Y_i[(\mu_i - r - \sigma_i'\theta)dt + \sigma_i' dW^Q] = Y_i\sigma_i' dW^Q,$$

which is Q-martingale increment because σ is assumed to be uniformly bounded. □

Corollary 4.18 (Q-Dynamics of the Assets)
Under an EMM Q, the dynamicss of asset i are given by

$$dS_i(t) = S_i(t)[r(t)dt + \sigma_i(t)' dW^Q(t)],$$

i.e. the drift μ does not matter under Q.

Theorem 4.19 (Arbitrage-free Pricing under Q and Completeness)
(i) Let Q be an EMM and let φ be an admissible portfolio strategy. Then $\{X^\varphi(t)/M(t)\}_{t\in[0,T]}$ is both a local Q-martingale and a Q-supermartingale implying that

$$\mathrm{E}_Q\left[\frac{X^\varphi(t)}{M(t)}\right] \leq X^\varphi(0) \qquad \text{for all } t \in [0,T].$$

Besides, the market is arbitrage-free. If additionally

$$\mathrm{E}_Q\left[\int_0^T \left(\varphi_i(s)\frac{S_i(s)}{M(s)}\right)^2 ds\right] < \infty \tag{24}$$

for every $i = 1, \ldots, I$, then $\{X^\varphi(t)/M(t)\}_{t\in[0,T]}$ is a Q-martingale.

(ii) Assume that $I = m$ ("# stocks = # sources of risk") and that σ is uniformly positive definite. Then (10) has a unique solution which is uniformly bounded implying that there exists a unique EMM Q. Besides, the market is complete.

Furthermore, let $C(T)$ be an $\mathcal{F}_T$-measurable contingent claim with

$$x := \mathrm{E}_Q[C(T)/M(T)] < \infty$$

Then the time-t price of the claim equals

$$C(t) = M(t)\,\mathrm{E}_Q\left[\left.\frac{C(T)}{M(T)}\right|\mathcal{F}_t\right],$$

i.e. the discounted price process $\{C(t)/M(t)\}_{t\in[0,T]}$ is a Q-martingale.

Proof. (i) Consider

$$dX^\varphi = \varphi' dS = Xrdt + \sum_{i=1}^I \varphi_i S_i \sigma_i' dW^Q.$$

Since $d(1/M) = -(1/M)rdt$,

$$d\Big(\frac{X}{M}\Big) = Xd\Big(\frac{1}{M}\Big) + \frac{1}{M}dX = \frac{1}{M}\sum_{i=1}^I S_i\varphi_i\sigma_i' dW^Q$$

which is a positive local Q-martingale and thus a Q-supermartingale. If (24) is satisfied, then it is a Q-martingale.

(ii) Since θ is unique, by Theorem 4.17, the EMM Q is unique. The claim price follows from Theorem 4.5 because $H = Z/M$. □

Remark.

- An EMM exists if, for instants, θ is bounded. This follows from Theorem 4.17.
- If φ is uniformly bounded, then (24) holds because μ, σ, and r are assumed to be uniformly bounded.
- If the market is **incomplete**, then there exist more than one market price of risk θ. Let Q and $\tilde{Q}$ be two corresponding EMM. If (24) holds, then for an admissible strategy φ we get

$$M(t)\,\mathrm{E}_Q\left[\left.\frac{X^\varphi(T)}{M(T)}\right|\mathcal{F}_t\right] = X^\varphi(t) = M(t)\,\mathrm{E}_{\tilde{Q}}\left[\left.\frac{X^\varphi(T)}{M(T)}\right|\mathcal{F}_t\right].$$

This implies that any EMM assigns the same value to a replicable claim, namly its replication costs. However, the prices of non-replicable claims may differ under different EMM. This is the reason why pricing in incomplete markets is so complicated and in some sense arbitrary.

Corollary 4.20 (Black-Scholes Again)
The price of the European call is given by

$$C(t) = S(t) \cdot N(d_1(t)) - K \cdot N(d_2(t)) \cdot e^{-r(T-t)}$$

with

$$\begin{aligned} d_1(t) &= \frac{\ln(S(t)/K) + (r + 0.5^2)(T-t)}{\sigma\sqrt{T-t}}, \\ d_2(t) &= d_1(t) - \sigma\sqrt{T-t}. \end{aligned}$$

Proof. By Theorem 4.19,

$$C(t) = \mathrm{E}_Q[\max\{S(T) - K; 0\}]\, e^{-r(T-t)},$$

where due to Corollary 4.18

$$dS(t) = S(t)[rdt + \sigma dW^Q(t)].$$

Thus $S(T)$ has the same distribution as $Y(T)$ in the proof of Proposition 4.3 and the result follows. □

Corollary 4.21 (Q-Dynamics of Claims)
Under the assumptions of Theorem 4.19, the dynamics of a replicable claim $C(T)$ are given by

$$dC(t) = C(t)[r(t)dt + \psi(t)'dW^Q(t)] \tag{25}$$

with a suitable progressively measurable process ψ.

Proof. By Theorem 4.19, C/M is a (local) Q-martingale and thus (25) follows from the martingale representation theorem. □

Proposition 4.22 (Black-Scholes PDE Again and Hedging)
Under the assumptions of Theorem 4.19, let $C(T)$ be a replicable claim with $C(t) = f(t, S(t))$ and $C(T) = h(S(T))$, where f is a $C^{1,2}$-function.

(i) The pricing function f satisfies the following multi-dimensional Black-Scholes PDE:

$$f_t + \sum_{i=1}^{I} f_i s_i r + 0.5 \sum_{i,k=1}^{I} \sum_{j=1}^{m} s_i s_k \sigma_{ij} \sigma_{kj} f_{ik} - rf = 0, \quad f(T, s) = h(s),\ s \in I\!R_+^I,$$

where f_i denotes the partial derivative w.r.t. the i-th asset price.

(ii) A replication strategy for $C(T)$ is given by

$$\varphi_i(t) = f_i(t, S(t)), \qquad i = 1, \dots, I,$$

where the remaining funds $C(t) - \sum_{i=1}^I \varphi_i(t)S_i(t)$ are invested in the money market account, i.e. $\varphi_M(t) = (C(t) - \sum_{i=1}^I \varphi_i(t)S_i(t))/M(t)$.

Proof. (i) Ito's formula yields

$$(26) \quad dC = \Big(f_t + \sum_{i=1}^I f_i S_i r + 0.5 \sum_{i,k=1}^I \sum_{j=1}^m f_{ik} S_i S_k \sigma_{ij}\sigma_{kj}\Big)dt + \sum_{i=1}^I \sum_{j=1}^m f_i S_i \sigma_{ij} dW_j^Q.$$

Since the representation of an Ito process is unique, comparing the drift with the drift of (25) yields the Black-Scholes PDE.

(ii) Since $C(T)$ is attainable, there exists a replication strategy φ with $X^\varphi(T) = C(T)$ and Q-dynamics

$$dX^\varphi = X^\varphi r dt + \sum_{i=1}^I \sum_{j=1}^m \varphi_i S_i \sigma_{ij} dW_j^Q,$$

where due to Theorem 4.19 (ii) the drift equals r.[27] Comparing the diffusion part with the diffusion part of (26) shows that one can choose $\varphi_i = f_i$. □

Remarks.

- The strategy in (ii) is said to be a **delta hedging strategy**.
- Under the conditions of Theorem 4.19 (ii), the delta hedging strategy is the unique replication strategy.
- In the Black-Scholes model, we get for the European call (see Exercise 29): $\varphi_S = N(d_1)$. Due to the put-call parity, the delta hedge for the put reads $\varphi_S = -N(-d_1)$.
- If $I > m$, i.e. ("# stocks > # sources of risk"), then in general there exist further replication strategies, i.e. the choice $\varphi_i = f_i$ in the proof of (ii) is not the only possible choice.

In Theorem 4.19, we have proved that the following implications hold:

1) Existence of EMM $\Longrightarrow$ No arbitrage

2) Uniqueness of EMM $\Longrightarrow$ Completeness

The following theorems show to which extend the converse is true.

Theorem 4.23 (Existence of a Market Price of Risk)
Assume that the market (6) is arbitrage-free. Then there exists a market price of risk, i.e. there exists a progressively measurable process θ solving the system (10).

Proof. Define $\phi_i := \varphi_i S_i$ (amount of money invested in asset i) leading to the following version of the wealth equation:

$$dX(t) = X(t)r(t)dt + \phi(t)'(\mu(t) - r(t)\mathbf{1})dt + \phi(t)'\sigma(t)dW(t).$$

[27] The representation of the diffusion part follows from rewriting $X\pi'\sigma dW^Q$ in terms of $\varphi = \pi X/S$.

Assume that there exists a self-financing strategy with

$$\phi(t)'\sigma(t) = 0 \quad \text{and} \quad \phi(t)'(\mu(t) - r(t)\mathbf{1}) \neq 0.$$

This strategy would be an arbitrage since it does not involve risk, but its wealth equation has a drift different from r. Thus in an arbitrage-free market every vector in the kernel $\mathrm{Ker}(\sigma(t)')$ should be orthogonal to the vector $\mu(t) - r(t)\mathbf{1}$. But the orthogonal complement of $\mathrm{Ker}(\sigma(t)')$ is $\mathrm{Range}(\sigma(t))$. Hence, $\mu(t) - r(t)\mathbf{1}$ should be in the range of $\sigma(t)$, i.e. there exists a $\theta(t)$ such that

$$\mu(t) - r(t)\mathbf{1} = \sigma(t)\theta(t)$$

The progressive measurability of θ is proved in Karatzas/Shreve (1999, pp. 12ff).□

Remark. In this theorem, nothing is said about whether the corresponding process Z is a martingale. Although a market price of risk exists, it is not clear if an EMM exists.

Theorem 4.24 (Uniqueness of EMM)
Assume that the market is complete and let Q be an EMM such that $\{X^\varphi(t)/M(t)\}_{t\in[0,T]}$ is a Q-martingale for all admissible strategies φ with $\mathrm{E}[X^\varphi(T)/M(T)] < \infty$. If $\tilde{Q}$ is another EMM, then $Q = \tilde{Q}$, i.e. Q is the unique EMM.

Proof. Fix $j \in \{1, \ldots, m\}$. Since the market is complete, there exists an admissible trading strategy φ_j such that $X^{\varphi_j}(T) = Y_j(T)$, where

$$Y_j(T) = \exp\Big(\int_0^T r(s)\,ds - 0.5T + W_j^Q(T)\Big).$$

From Theorem 4.19, we know that the process X^{φ_j}/M is a local martingale under Q and $\tilde{Q}$. By assumption, X^{φ_j}/M is even a Q-martingale implying

$$X^{\varphi_j}(t) = M(t)\mathrm{E}_Q\left[\frac{X^{\varphi_j}(T)}{M(T)}\bigg|\mathcal{F}_t\right] = M(t)\mathrm{E}_Q\left[\frac{Y_j(T)}{M(T)}\bigg|\mathcal{F}_t\right] = Y_j(t) \quad P-a.s.,$$

i.e. Y_j is a modification of X^{φ_j}. By Proposition 3.2, X^{φ_j} and Y_j are indistinguishable. Hence, Y_j/M is a local martingale under Q and $\tilde{Q}$. By Theorem 4.17, both measures are characterized by market prices of risk θ and $\tilde{\theta}$ such that

$$\begin{aligned} dW_j^Q(t) &= dW_j(t) + \theta(t)dt, \\ dW_j^{\tilde{Q}}(t) &= dW_j(t) + \tilde{\theta}(t)dt. \end{aligned}$$

Therefore,

$$\frac{Y_j(t)}{M(t)} = 1 + \int_0^t \frac{Y_j(s)}{M(s)}\,dW_j^Q(s) = 1 + \int_0^t \frac{Y_j(s)}{M(s)}\,dW_j^{\tilde{Q}}(s) + \int_0^t \frac{Y_j(s)}{M(s)}(\theta_j(s) - \tilde{\theta}_j(s))\,ds.$$

Since Y_j/M is a Brownian local $\tilde{Q}$-martingale, the process $Y_j(\theta_j - \tilde{\theta}_j)/M$ needs to be indistinguishable from zero.[28] Since $Y_j/M > 0$, the market prices of risk θ_j and $\tilde{\theta}_j$ are indistinguishable. Since this holds for every j, we get $Q = \tilde{Q}$. □

Remarks.

[28]Recall that the representation of Ito processes is unique up to indistinguishability.

- Since $\{X^\varphi(t)/M(t)\}_{t\in[0,T]}$ is a Q-supermartingale, this process is a Q-martingale if, for instants,
$$\mathrm{E}_Q\left[\frac{X^\varphi(T)}{M(T)}\right] = X^\varphi(0).$$
- Obviously it is sufficient to check if $\{X^{\varphi_j}(t)/M(t)\}_{t\in[0,T]}$ is a Q-martingale for every $j = 1, \ldots, m$.

5 Continuous-time Portfolio Optimization

5.1 Motivation

In the theory of **option pricing** the following problem is tackled:

Given: Payoff $C(T)$

Wanted: Today's price $C(0)$ and hedging strategy φ

In the theory of **portfolio optimization** the converse problem is considered:

Given: Initial wealth $X(0)$

Wanted: Optimal final wealth $X^*(T)$ and optimal strategy φ^*

Question: What is "optimal"?

1st Idea. Maximze expected final wealth, i.e. $\max_\varphi \mathrm{E}[X^\varphi(T)]$.

Problem: Petersburg Paradox. Consider a game where a coin is flipped. If in the n-th round, the first time head occurs, then the player gets 2^{n-1} Euros, i.e. the expected value is

$$\frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{8}\cdot 4 + \ldots + \frac{1}{2^n}\cdot 2^{n-1} + \ldots = \sum_{n=0}^{\infty} \frac{1}{2} = \infty.$$

Hence, investors who maximize expected final wealth should be willing to pay ∞ Euros to play this game! Therefore, maximization of expected final wealth is usually not a reasonable criterion.

Question: What is the reason for this result?

Answer: People are risk averse, i.e. if they can choose between
- 50 Euros or
- 100 Euros with probability 0.5 and 0 Euros with probability 0.5,

they choose the 50 Euros.

Bernoulli's idea. Measure payoffs in terms of a utility function $U(x) = \ln(x)$, i.e.

$$\begin{aligned}&\frac{1}{2}\cdot\ln(1) + \frac{1}{4}\cdot\ln(2) + \frac{1}{8}\cdot\ln(4) + \ldots + \frac{1}{2^n}\cdot\ln(2^{n-1}) + \ldots\\ &= \sum_{n=1}^{\infty}\frac{1}{2^n}\cdot\ln(2^{n-1}) = \ln(2)\sum_{n=1}^{\infty}\frac{n-1}{2^n} = \ln(2).\end{aligned}$$

Hence, an investor following this criterion would be willing to pay at most 2 Euros to play this game.

This idea can be generalized: For a given initial wealth $x_0 > 0$, the investor maximizes the following utility functional

$$\sup_{\pi \in \mathcal{A}} \mathrm{E}[U(X^\pi(T))], \tag{27}$$

where $\mathcal{A} := \{\pi : \pi \text{ admissible and } \mathrm{E}[U(X^\pi(T))^-] < \infty\}$ and

$$dX^\pi(t) = X^\pi(t)\Big[(r(t) + \pi(t)'(\mu(t) - r(t)\mathbf{1}))dt + \pi(t)'\sigma dW(t)\Big], \quad X^\pi(0) = x_0,$$

and the utility function U has the following properties:

Definition 5.1 (Utility Function)
A strictly concave, strictly increasing, and continuously differentiable function $U : (0,\infty) \to I\!R$ is said to be a utility function if

$$U'(0) := \lim_{x \searrow 0} U'(x) = \infty, \qquad U'(\infty) := \lim_{x \nearrow \infty} U'(x) = 0.$$

Remarks.

- The investor's **greed** is reflected by $U' > 0$, i.e. loosely speaking: "More is better than less".
- The investor's **risk aversion** is reflected by the concavity of U which implies that for any non-constant random variable Z, by Jensen's inequality,

  $$\mathrm{E}[U(Z)] < U(\mathrm{E}[Z]).$$

  i.e. loosely speaking: "A bird in the hand is worth two in the bush".

There exist two approaches to solve the dynamic optimization problem (27):

1. Martingale approach: Theorem 4.5 ensures that in a complete market any claim can be replicated.
$\Longrightarrow$ Determine the optimal wealth via a static optimization problem and interpret this wealth as a contingent claim

2. Optimal control approach: We know that expectations can be represented as solutions to PDEs (Feynman-Kac theorem)
$\Longrightarrow$ Represent the functional (27) as PDE and solve the PDE

Standing Assumption. We consider a Brownian market (6) with bounded coefficients.

5.2 Martingale Approach

Assumption in the Martingale Approach: The market is complete.

The martingale approach consists of a **three-step procedure**:

1st step. Compute the optimal wealth $X^*(T)$ by solving a static optimization problem.

2nd step. Prove that the static optimization problem is equivalent to the dynamic optimization problem (27)

3rd step. Determine the replication strategy for $X^*(T)$ which, in the context of portfolio optimization, is said to be the optimal portfolio strategy π^*.

1st step. Consider the **constrained static optimization problem**

$$\sup_{\mathcal{X}\in\mathcal{C}} \mathrm{E}[U(\mathcal{X})],$$

subject to the so-called budget constraint (BC)

$$\mathrm{E}[H(T)\mathcal{X}] \le x_0$$

and $\mathcal{C} := \{C(T) \ge 0 : C(T) \text{ is } \mathcal{F}_T\text{-measurable and } \mathrm{E}[U(C(T))^-] < \infty\}$.

The BC states that the present value of terminal wealth cannot exceed the initial wealth x_0 (see Theorem 4.5 (i)).

Heuristic solution. Pointwise Lagrangian

$$L(\mathcal{X},\lambda) := U(\mathcal{X}) + \lambda(x_0 - H(T)\mathcal{X}).$$

First-order condition (FOC)

$$U'(\mathcal{X}) - \lambda H(T) = 0,$$

i.e. $\mathcal{X}^* = V(\lambda H(T))$, where $V := (U')^{-1} : (0,\infty) \to (0,\infty)$ is strictly decreasing.

Since the investor is greedy, the BC is satisfied as equality and we can choose λ^* such that

$$\Gamma(\lambda) := \mathrm{E}[H(T)V(\lambda H(T))] = x_0 \tag{28}$$

leading to $\lambda^* = \Lambda(x_0) := \Gamma^{-1}(x_0)$.

2nd step. Verify that the heuristic solution is indeed the optimal solution.

Lemma 5.1 (Properties of Γ)
Assume $\Gamma(\lambda) < \infty$ for all $\lambda > 0$. Then Γ is

(i) strictly decreasing,

(ii) continuous on $(0,\infty)$,

(iii) $\Gamma(0) := \lim_{\lambda\searrow 0}\Gamma(\lambda) = \infty$, and $\Gamma(\infty) := \lim_{\lambda\nearrow\infty}\Gamma(\lambda) = 0$.

Proof. (i) V is strictly decreasing on $(0,\infty)$. Since $H(t) > 0$ for every $t \in [0,T]$, the function Γ is also strictly decreasing.

(ii) follows from the continuity of H and V as well as the dominated convergence theorem.[29]

[29] To check continuity at λ_0, choose some $\lambda_1 < \lambda_0$. Then due to (i) we can use $\Gamma(\lambda_1) < \infty$ as a majorant.

(iii) Since

$$\lim_{y \searrow 0} V(y) = \infty, \qquad \lim_{y \nearrow \infty} V(y) = 0,$$

and V is strictly decreasing, the first result follows from the monotonous convergence theorem and the second from the dominated convergence theorem. □

Remark. Hence, $\Lambda(x) = \Gamma^{-1}(x)$ exists on $(0, \infty)$ with $\Lambda \geq 0$ and

$$\Lambda(0) := \lim_{x \searrow 0} \Lambda(x) = \infty, \qquad \Lambda(\infty) := \lim_{x \nearrow \infty} \Lambda(x) = 0.$$

Therefore, one can always find a λ such that (28) is satisfied as equality.

Lemma 5.2
For a utility function U with $U^{-1} = V$ we get

$$U(V(y)) \geq U(x) + y(V(y) - x) \quad \text{for } 0 < x, y < \infty.$$

Proof. By concavity of U, Taylor approximation leads to $U(x_2) \leq U(x_1) + U'(x_1)(x_2 - x_1)$ and thus $U(x_1) \geq U(x_2) + U'(x_1)(x_1 - x_2)$ implying

$$U(V(y)) \geq U(x) + U'(V(y))(V(y) - x) = U(x) + y(V(y) - x).$$

□

Theorem 5.3 (Optimal Final Wealth)
Assume $x_0 > 0$ and $\Gamma(\lambda) < \infty$ for all $\lambda > 0$. Then the optimal final wealth reads

$$\mathcal{X}^* = V(\Lambda(x_0) \cdot H(T))$$

and there exists an admissible optimal strategy $\pi^* \in \mathcal{A}$ such that $X^{\pi^*}(0) = x_0$ and $X^{\pi^*}(T) = \mathcal{X}^*$, P-a.s.

Proof. By definition of $\Lambda(x_0)$, we get $\mathrm{E}[H(T)\mathcal{X}^*] = x_0$. Besides, $\mathcal{X}^* > 0$ because $V > 0$. Hence, $\mathcal{X}^*$ satisfies all constraints and, by Theorem 4.5 (ii), there exists a strategy π^* replicating $\mathcal{X}^*$.

We now verify $\mathrm{E}[U(\mathcal{X}^*)^-] < \infty$, i.e. $\pi^* \in \mathcal{A}$. Due to Lemma 5.2,

$$U(\mathcal{X}^*) \geq U(1) + \Lambda(x_0)H(T)(\mathcal{X}^* - 1).$$

Therefore,[30]

$$\mathrm{E}[U(\mathcal{X}^*)^-] \leq |U(1)| + \Lambda(x_0)\,\mathrm{E}[H(T)(\mathcal{X}^* + 1)] \leq |U(1)| + \Lambda(x_0)\,(x_0 + \mathrm{E}[H(T)]) < \infty.$$

Finally, we show that π^* and $\mathcal{X}^*$ are indeed optimal. Consider some $\pi \in \mathcal{A}$. Due to Lemma 5.2,

$$U(\mathcal{X}^*) \geq U(X^\pi(T)) + \Lambda(x_0)H(T)(\mathcal{X}^* - X^\pi(T))$$

[30] $a \geq b$ implies $a^- \leq b^- \leq |b|$.

and thus

$$\begin{aligned}
\mathrm{E}[U(\mathcal{X}^*)] &\geq \mathrm{E}[U(X^\pi(T))] + \Lambda(x_0)\mathrm{E}[H(T)(\mathcal{X}^* - X^\pi(T))] \\
&= \mathrm{E}[U(X^\pi(T))] + \Lambda(x_0)\underbrace{(x_0 - \mathrm{E}[H(T)(X^\pi(T))])}_{\overset{BC}{\geq}\, 0}
\end{aligned}$$

which gives the desired result. $\square$

3rd step. To determine the optimal strategy π^*, one can apply a similar method as in Proposition 4.22 (ii). This is demonstrated in the following example.

Example "Power Utility with Constant Coefficients".
We make the following assumptions:
- Power utility function $U(x) = \frac{1}{\gamma}x^\gamma$
- $I = m = 1$ ("one stock, one Brownian motion")
- The coefficients r, μ, σ are constant.
- $\sigma > 0$, i.e. the market is complete.

Since $V(y) = (U')^{-1}(y) = y^{\frac{1}{\gamma-1}}$, the optimal final wealth reads:

$$\mathcal{X}^* = V(\Lambda(x_0)H(T)) = (\Lambda(x_0)H(T))^{\frac{1}{\gamma-1}},$$

where

$$H(T)^{\frac{1}{\gamma-1}} = \exp\Big(\tfrac{1}{1-\gamma}rT + 0.5\tfrac{1}{1-\gamma}\theta^2 T + \tfrac{1}{1-\gamma}\theta W(T)\Big) \tag{29}$$

and $\theta = (\mu - r)/\sigma$. On the other hand, the final wealth for a strategy π is given by

$$X^\pi(T) = x_0 \exp\Big(rT + \int_0^T (\mu - r)\pi_s - 0.5\sigma^2\pi_s^2\, ds + \int_0^t \sigma\pi_s\, dW_s\Big). \tag{30}$$

By comparing the diffusion parts of (29) and (30) we guess the optimal strategy:

$$\pi^*(t) = \frac{1}{1-\gamma}\frac{\mu - r}{\sigma^2}$$

implying

$$X^{\pi^*}(T) = \exp\Big(rT + \tfrac{1}{1-\gamma}\theta^2 T - 0.5\tfrac{1}{(1-\gamma)^2}\theta^2 T + \tfrac{1}{1-\gamma}\theta W(T)\Big).$$

Since $\Gamma(\lambda) = \mathrm{E}[H(T)V(\lambda H(T))] = \lambda^{\frac{1}{\gamma-1}}\,\mathrm{E}[H(T)^{\frac{\gamma}{\gamma-1}}] = x_0$, we get

$$\Lambda(x_0)^{\frac{1}{\gamma-1}} = \Gamma^{-1}(x_0) = x_0 \exp\Big(-\tfrac{\gamma}{1-\gamma}rT - 0.5\tfrac{\gamma}{1-\gamma}\theta^2 T - 0.5\big(\tfrac{\gamma}{1-\gamma}\big)^2\theta^2 T\Big).$$

Therefore,

$$\mathcal{X}^* = (\Lambda(x_0)H(T))^{\frac{1}{\gamma-1}} = x_0 \exp\Big(rT + 0.5\theta^2 T - 0.5\big(\tfrac{\gamma}{1-\gamma}\big)^2\theta^2 T + \tfrac{1}{1-\gamma}\theta W(T)\Big).$$

It is easy to check that

$$\tfrac{1}{1-\gamma} - 0.5\tfrac{1}{(1-\gamma)^2} = 0.5 - 0.5\big(\tfrac{\gamma}{1-\gamma}\big)^2.$$

Hence, $\mathcal{X}^* = X^{\pi^*}(T)$ and due to completeness π^* is the unique optimal portfolio strategy.

5.3 Stochastic Optimal Control Approach

Assumption. The coefficients r, μ, σ are constant and $I = m = 1$.

Define the value function

$$G(t,x) := \sup_{\pi \in \mathcal{A}} \mathrm{E}^{t,x}[U(X^\pi(T))],$$

where $\mathrm{E}^{t,x}[U(X^\pi(T))] := \mathrm{E}[U(X^\pi(T))|X(t) = x]$. Note that $G(T,x) = U(x)$.

5.3.1 Heuristic Derivation of the HJB

We start with a **heuristic derivation** of the so-called Hamiltion-Jacobi-Bellman equation (HJB) which is a PDE for G.

Assume that there exists an optimal portfolio strategy π^*. We now carry out the following **procedure**:

(i) For given $(t,x) \in [0,T] \times (0,\infty)$, we consider the following strategies over $[t,T]$:
- Strategy I: Use the optimal strategy π^*.
- Strategy II: For an arbitrary π use the strategy

$$\hat{\pi} = \begin{cases} \pi & \text{on} \quad [t, t+\Delta] \\ \pi^* & \text{on} \quad (t+\Delta, T] \end{cases}$$

(ii) Compute $\mathrm{E}^{t,x}[U(X^{\pi^*}(T))]$ and $\mathrm{E}^{t,x}[U(X^{\hat{\pi}}(T))]$.

(iii) Using the fact that π^* is at least as good as $\hat{\pi}$ and taking the limit $\Delta \to 0$ yields the HJB.

We set $X^* := X^{\pi^*}$ and $\hat{X} := X^{\hat{\pi}}$.

ad (ii).

Strategy I. By assumption, $\mathrm{E}^{t,x}[U(X^*(T))] = G(t,x)$.

Strategy II.

$$\begin{aligned} \mathrm{E}^{t,x}[U(\hat{X}(T))] &= \mathrm{E}^{t,x}[\mathrm{E}^{t+\Delta,\hat{X}(t+\Delta)}[U(\hat{X}(T))]] = \mathrm{E}^{t,x}[\mathrm{E}^{t+\Delta,\hat{X}(t+\Delta)}[U(X^*(T))]] \\ &= \mathrm{E}^{t,x}[G(t+\Delta,\hat{X}(t+\Delta))] = \mathrm{E}^{t,x}[G(t+\Delta, X^\pi(t+\Delta))], \end{aligned}$$

for any admissible π.

ad (iii). By definition of the strategies,

$$\mathrm{E}^{t,x}[G(t+\Delta, X^\pi(t+\Delta))] \le G(t,x), \tag{31}$$

where we have equality if $\pi = \pi^*$. Given that $G \in C^{1,2}$, Ito's formula yields

$$G(t+\Delta, X^\pi(t+\Delta)) = G(t,x) + \int_t^{t+\Delta} G_t(s, X_s^\pi)ds + \int_t^{t+\Delta} G_x(s, X_s^\pi)dX_s^\pi$$

$$
\begin{aligned}
&\quad +0.5\int_t^{t+\Delta} G_{xx}(s,X_s^\pi)d<X^\pi>_s \\
&= G(t,x)+\int_t^{t+\Delta}\Big(G_t(s,X_s^\pi)+G_x(s,X_s^\pi)X_s^\pi(r+\pi_s(\mu-r)) \\
&\qquad +0.5G_{xx}(s,X_s^\pi)(X_s^\pi)^2\pi_s^2\sigma^2\Big)\,ds \\
&\quad +\int_t^{t+\Delta} G_x(s,X_s^\pi)X_s^\pi\pi_s\sigma\,dW_s
\end{aligned}
$$

Given $\mathrm{E}^{t,x}[\int_t^{t+\Delta} G_x(s,X_s^\pi)X_s^\pi\pi_s\sigma\,dW_s]=0$, we can rewrite (31):

$$
\begin{aligned}
\mathrm{E}^{t,x}\Big[\int_t^{t+\Delta}&\Big(G_t(s,X_s^\pi)+G_x(s,X_s^\pi)X_s^\pi(r+\pi_s(\mu-r)) \\
&+0.5G_{xx}(s,X_s^\pi)(X_s^\pi)^2\pi_s^2\sigma^2\Big)\,ds\Big]\le 0.
\end{aligned}
$$

Dividing by Δ and taking the limit $\Delta\to 0$ we get (Note: $X^\pi(t)=x$)

$$G_t(t,x)+G_x(t,x)x(r+\pi_t(\mu-r))+0.5G_{xx}(t,x)x^2\pi_t^2\sigma^2\le 0$$

for any admissible π. As above, we have equality for $\pi=\pi^*$. Therefore, we get the **HJB**

$$\sup_\pi\left\{G_t(t,x)+G_x(t,x)x(r+\pi(\mu-r))+0.5G_{xx}(t,x)x^2\pi^2\sigma^2\right\}=0.$$

with final condition $G(T,x)=U(x)$.

Remarks.

- Since (t,x) was arbitrarily chosen, the HJB holds for all $(t,x)\in[0,T]\times(0,\infty)$.
- Note that we take the supremum over a number π and not over a process.

5.3.2 Solving the HJB

We consider the HJB for an investor with power utility $U(x)=\frac{1}{\gamma}x^\gamma$:

$$\sup_\pi\left\{G_t(t,x)+G_x(t,x)x(r+\pi(\mu-r))+0.5G_{xx}(t,x)x^2\pi^2\sigma^2\right\}=0.$$

with final condition $G(T,x)=\frac{1}{\gamma}x^\gamma$. Pointwise maximization over π yields the following FOC

$$G_x(t,x)x(\mu-r)+G_{xx}(t,x)x^2\pi(t)\sigma^2=0 \tag{32}$$

implying

$$\pi^*(t)=-\frac{\mu-r}{\sigma^2}\frac{G_x(t,x)}{xG_{xx}(t,x)}.$$

Since the derivative of (32) w.r.t. π is $G_xx(t,x)x^2\sigma^2$, the candidate π^* is the global maximum if $G_{xx}<0$.

Substituting π^* into the HJB yields a PDE for G:

$$G_t(t,x)+rxG_x(t,x)-0.5\theta^2\frac{G_x^2(t,x)}{G_{xx}(t,x)}=0$$

with $G(T,x) = \frac{1}{\gamma}x^\gamma$. We try the separation $G(t,x) = \frac{1}{\gamma}x^\gamma f(t)^{1-\gamma}$ for some function f with $f(T) = 1$ and obtain

$$f' = \underbrace{-\tfrac{\gamma}{1-\gamma}(r + 0.5\tfrac{1}{1-\gamma}\theta^2)}_{=:K} f.$$

This is an ODE for f with the solution $f(t) = \exp(-K(T-t))$. Hence, we get the following candidates for the value function and the optimal strategy:

$$\begin{aligned} G(t,x) &= \frac{1}{\gamma}x^\gamma \exp\big(-(1-\gamma)K(T-t)\big), \\ \pi^*(t) &= \frac{1}{1-\gamma}\frac{\mu - r}{\sigma^2}. \end{aligned} \tag{33}$$

Remark. Our candidate G is indeed a $C^{1,2}$-function. Since $\gamma < 1$, we also get $G_{xx}(t,x) = (\gamma-1)x^{\gamma-1}f(t) < 0$.

5.3.3 Verification

In this section, we show that our heuristic derivation of the HJB has lead to the correct result.

Proposition 5.4 (Verification Result)
Assume that the coefficients are constant and that the investor has a power utility function. Then candidates (33) are the value function and the optimal strategy of the portfolio problem (27), where for $\gamma < 0$ only bounded admissible strategies are considered.

Proof. Let $\pi \in \mathcal{A}$. Since $G \in C^{1,2}$, Ito's formula yields

$$\begin{aligned} G(T, X^\pi(T)) = G(t,x) &+ \int_t^T G_t(s, X_s^\pi)ds + \int_t^T G_x(s, X_s^\pi)dX_s^\pi \\ &+ 0.5\int_t^T G_{xx}(s, X_s^\pi)d < X^\pi >_s \\ = G(t,x) + \int_t^T &\Big(G_t(s, X_s^\pi) + G_x(s, X_s^\pi)X_s^\pi(r + \pi_s(\mu - r)) \\ +0.5G_{xx}&(s, X_s^\pi)(X_s^\pi)^2\pi_s^2\sigma^2\Big)\, ds + \int_t^T G_x(s, X_s^\pi)X_s^\pi\pi_s\sigma\, dW_s \\ \overset{(*)}{\leq} G(t,x) + &\int_t^T G_x(s, X_s^\pi)X_s^\pi\pi_s\sigma\, dW_s, \end{aligned}$$

where $(*)$ holds because G satisfies the HJB.

The right-hand side is a local martingale in T. For $\gamma > 0$ we have $G > 0$ and thus the right-hand side is even a supermartingale implying

$$\mathrm{E}^{t,x}[\tfrac{1}{\gamma}(X^\pi(T))^\gamma] \leq G(t,x), \tag{34}$$

where we use $G(T,x) = \frac{1}{\gamma}x^\gamma$. Besides, since $\mathrm{E}^{t,x}[\int_t^T G_x(s, X_s^*)X_s^*\pi_s^*\sigma\, dW_s] = 0$ (exercise) and for π^*

$$G_t(s, X_s^*) + G_x(s, X_s^*)X_s^*(r + \pi_s^*(\mu - r)) + 0.5G_{xx}(s, X_s^*)(X_s^*)^2(\pi_s^*)^2\sigma^2 = 0,$$

we obtain

$$\mathrm{E}^{t,x}[\tfrac{1}{\gamma}(X^*(T))^\gamma] = G(t,x),$$

i.e. G is the value function and π^* is the optimal solution.

Since for a bounded strategy π we get $\mathrm{E}[\int_t^T G_x(s, X_s^\pi) X_s^\pi \pi_s \sigma \, dW_s] = 0$, the relation (34) also holds for $\gamma < 0$ and bounded admissible strategies. □

Remarks.
- For $\gamma < 0$ the boundedness assumption of π can be weakened.
- Note that we have not explicitly used the fact that the market is complete.

5.4 Intermediate Consumption

So far the investor has maximized expected utility from terminal wealth only. We now consider a situation where the investor maximizes

- expected utility from terminal wealth as well as
- expected utility from intermediate consumption.

Assumption. At time t, the investor consumes at a rate $c(t) \geq 0$, i.e. over the interval $[0,T]$ he consumes $\int_0^T c(t)\,dt$.

Remark. Hence, in our model consumption can be interpreted as a continuous stream of dividends (See Exercise 40).

To formalize the problem we need to generalize the concept of admissibility:

Definition 5.2 (Consumption Process, Self-financing, Admissible)
(i) A progressively measurable, real-valued process $\{c_t\}_{t\in[0,T]}$ with $c \geq 0$ and $\int_0^T c(t)\,dt < \infty$ $P-a.s.$ is said to be a consumption process.

(ii) Let φ be a trading strategy (see Definition 4.1) and c be a consumption process. If

$$dX(t) = \varphi(t)' dS(t) - c(t)dt$$

holds, then the pair (φ, c) is said to be self-financing.

(iii) A self-financing (φ, c) is said to be admissible if $X(t) \geq 0$ for all $t \geq 0$.

Hence, the investor's goal function reads

$$\sup_{(\pi,c)\in\mathcal{A}} \mathrm{E}\Big[\int_0^T U_1(t, c(t))\,dt + U_2(X^{\pi,c}(T))\Big], \tag{35}$$

where U_1 is a utility function for fixed $t \in [0,T]$,

$$\mathcal{A} := \Big\{(\pi, c) : (\pi, c) \text{ admissible and } \mathrm{E}\Big[\int_0^T U_1(t, c(t))^- \,dt + U_2(X^{\pi,c}(T))^-\Big] < \infty\Big\}$$

and

$$
\begin{aligned}
dX^{\pi,c}(t) &= X^{\pi,c}(t)\Big[(r(t)+\pi(t)'(\mu(t)-r(t)\mathbf{1}))dt+\pi(t)'\sigma dW(t)\Big]-c(t)dt,\\
X^{\pi,c}(0) &= x_0.
\end{aligned}
$$

We set $V_1 := (U_1')^{-1}$ and $V_2 := (U_2')^{-1}$ and define

$$
\Gamma(\lambda) := \mathrm{E}\Big[\int_0^T H(t)V_1(t,\lambda H(t))\,dt + H(T)V_2(\lambda H(T))\Big]
$$

It is straightforward to generalize the result of Section 5.2.

Theorem 5.5 (Optimal Consumption and Optimal Final Wealth)
Assume $x_0 > 0$ and $\Gamma(\lambda) < \infty$ for all $\lambda > 0$. Then the optimal consumption and the optimal final wealth reads

$$
\begin{aligned}
c^*(t) &= V_1(t,\Lambda(x_0)\cdot H(t))\\
\mathcal{X}^* &= V_2(\Lambda(x_0)\cdot H(T))
\end{aligned}
$$

and there exists a portfolio strategy π^* such that $(\pi^*,c^*)\in\mathcal{A}$ and $X^{\pi^*,c^*}(T)=x_0$ and $X^{\pi^*,c^*}(T)=\mathcal{X}^*$, P-a.s.

We will present the solution by using the **stochastic control approach**. The value function now reads

$$
G(t,x) := \sup_{(\pi,c)\in\mathcal{A}} \mathrm{E}^{t,x}\Big[\int_t^T U_1(s,c(s))\,ds + U_2(X^{\pi,c}(T))\Big].
$$

As in Section 5.3 we make the following

Assumption. The coefficients r, μ, σ are constant and $I=m=1$.

Assume that there exists an optimal strategy (π^*,c^*). We now carry out the **procedure** of Section 5.3:

(i) For given $(t,x)\in[0,T]\times(0,\infty)$, we consider the following strategies over $[t,T]$:

- Strategy I: Use the optimal strategy (π^*,c^*).
- Strategy II: For an arbitrary (π,c) use the strategy

$$
(\hat\pi,\hat c) = \begin{cases} (\pi,c) & \text{on} \quad [t,t+\Delta] \\ (\pi^*,c^*) & \text{on} \quad (t+\Delta,T] \end{cases}
$$

(ii) Compute $\mathrm{E}^{t,x}[\int_t^T U_1(s,c^*(s))\,ds + U_2(X^{\pi^*,c^*}(T))]$ and $\mathrm{E}^{t,x}[\int_t^T U_1(s,\hat c(s))\,ds + U_2(X^{\hat\pi,\hat c}(T))]$.

(iii) Using the fact that (π^*,c^*) is at least as good as $(\hat\pi,\hat c)$ and taking the limit $\Delta\to 0$ yields the HJB.

We set $X^* := X^{\pi^*,c^*}$ and $\hat X := X^{\hat\pi,\hat c}$.

ad (ii).

Strategy I. By assumption, $\mathrm{E}^{t,x}[\int_t^T U_1(s,c^*(s))\,ds + U_2(X^*(T))] = G(t,x)$.

Strategy II.

$$
\begin{aligned}
&\mathrm{E}^{t,x}\Big[\int_t^T U_1(s,\hat{c}(s))\,ds + U_2(\hat{X}(T))\Big] \\
&\quad = \mathrm{E}^{t,x}\Big[\mathrm{E}^{t+\Delta,\hat{X}(t+\Delta)}\Big[\int_t^T U_1(s,\hat{c}(s))\,ds + U_2(\hat{X}(T))\Big]\Big] \\
&\quad = \mathrm{E}^{t,x}\Big[\int_t^{t+\Delta} U_1(s,\hat{c}(s))\,ds + \mathrm{E}^{t+\Delta,\hat{X}(t+\Delta)}\Big[\int_{t+\Delta}^T U_1(s,\hat{c}(s))\,ds + U_2(\hat{X}(T))\Big]\Big] \\
&\quad = \mathrm{E}^{t,x}\Big[\int_t^{t+\Delta} U_1(s,c(s))\,ds + \mathrm{E}^{t+\Delta,\hat{X}(t+\Delta)}\Big[\int_{t+\Delta}^T U_1(s,c^*(s))\,ds + U_2(X^*(T))\Big]\Big] \\
&\quad = \mathrm{E}^{t,x}\Big[\int_t^{t+\Delta} U_1(s,c(s))\,ds + G(t+\Delta, X^{\pi,c}(t+\Delta))\Big]
\end{aligned}
$$

for any admissible (π, c).

ad (iii). By definition of the strategies,

$$
\mathrm{E}^{t,x}\Big[\int_t^{t+\Delta} U_1(s,c(s))\,ds + G(t+\Delta, X^{\pi,c}(t+\Delta))\Big] \le G(t,x), \tag{36}
$$

where we have equality if $(\pi,c) = (\pi^*, c^*)$. To shorten notations, we set $X := X^{\pi,c}$. Given that $G \in C^{1,2}$, Ito's formula yields

$$
\begin{aligned}
G(t+\Delta, X_{t+\Delta}) &= G(t,x) + \int_t^{t+\Delta} G_t(s,X_s)ds + \int_t^{t+\Delta} G_x(s,X_s)dX_s \\
&\qquad\qquad +0.5\int_t^{t+\Delta} G_{xx}(s,X_s)d<X>_s \\
&= G(t,x) + \int_t^{t+\Delta}\Big(G_t(s,X_s) + G_x(s,X_s)X_s(r+\pi_s(\mu-r)) - G_x(s,X_s)c_s \\
&\qquad\qquad +0.5G_{xx}(s,X_s)(X_s)^2\pi_s^2\sigma^2\Big)\,ds \\
&\quad + \int_t^{t+\Delta} G_x(s,X_s)X_s\pi_s\sigma\,dW_s
\end{aligned}
$$

Given $\mathrm{E}^{t,x}[\int_t^{t+\Delta} G_x(s,X_s)X_s\pi_s\sigma\,dW_s] = 0$, we can rewrite (36):

$$
\begin{aligned}
&\mathrm{E}^{t,x}\Big[\int_t^{t+\Delta} U_1(s,c_s)\,ds + \int_t^{t+\Delta}\Big(G_t(s,X_s) + G_x(s,X_s)X_s(r+\pi_s(\mu-r)) \\
&\qquad -G_x(s,X_s)c_s + 0.5G_{xx}(s,X_s)(X_s)^2\pi_s^2\sigma^2\Big)\,ds\Big] \le 0.
\end{aligned}
$$

Dividing by Δ and taking the limit $\Delta \to 0$ we get (Note: $X(t) = x$)

$$
U_1(t,c_t) + G_t(t,x) + G_x(t,x)x(r+\pi_t(\mu-r)) - G_x(t,x)c_t + 0.5G_{xx}(t,x)x^2\pi_t^2\sigma^2 \le 0
$$

for any admissible (π, c). As above, we have equality for $(\pi,c) = (\pi^*, c^*)$. Therefore, we get the **HJB**

$$
\begin{aligned}
\sup_{\pi,c}\Big\{&U_1(t,c) + G_t(t,x) + G_x(t,x)x(r+\pi(\mu-r)) - G_x(t,x)c \\
&+0.5G_{xx}(t,x)x^2\pi^2\sigma^2\Big\} = 0
\end{aligned} \tag{37}
$$

with final condition $G(T,x) = U_2(x)$.

Remarks.

- Since (t,x) was arbitrarily chosen, the HJB holds for all $(t,x) \in [0,T] \times (0,\infty)$.
- Note that we take the supremum over a number (π, c) and not over a process.

Example "Power Utility". We consider the following power utility functions

$$\begin{aligned} U_1(t,c) &= e^{-\alpha t}\frac{1}{\gamma}c^{\gamma}, \\ U_2(x) &= e^{-\alpha T}\frac{1}{\gamma}x^{\gamma}, \end{aligned}$$

where $\alpha \geq 0$ resembles the investor's time preferences. The HJB reads

$$\sup_{\pi,c}\Big\{e^{-\alpha t}\tfrac{1}{\gamma}c^{\gamma} + G_t(t,x) + G_x(t,x)x(r+\pi(\mu-r)) - G_x(t,x)c \\ +0.5G_{xx}(t,x)x^2\pi^2\sigma^2\Big\} = 0.$$

with final condition $G(T,x) = e^{-\alpha T}\frac{1}{\gamma}x^{\gamma}$. The FOCs read

$$\begin{aligned} G_x(t,x)x(\mu-r) + G_{xx}(t,x)x^2\pi(t)\sigma^2 &= 0 \\ e^{-\alpha t}c^{\gamma-1} - G_x(t,x) &= 0 \end{aligned}$$

implying

$$\begin{aligned} \pi^*(t) &= -\frac{\mu-r}{\sigma^2}\frac{G_x(t,x)}{xG_{xx}(t,x)} \\ c^*(t) &= e^{\frac{\alpha}{\gamma-1}t}G_x^{\frac{1}{\gamma-1}}(t,x) \end{aligned}$$

Substituting (π^*, c^*) into the HJB yields

$$\tfrac{1-\gamma}{\gamma}e^{\frac{\alpha}{\gamma-1}t}G_x^{\frac{\gamma}{\gamma-1}}(t,x) + G_t(t,x) + rxG_x(t,x) - 0.5\theta^2\frac{G_x^2(t,x)}{G_{xx}(t,x)} = 0.$$

We try the separation $G(t,x) = \frac{1}{\gamma}x^{\gamma}f^{\beta}(t)$ for some function f with $f(T) = 1$ and some constant β. Simplifying yields

$$\tfrac{1-\gamma}{\gamma}e^{-\frac{\alpha}{1-\gamma}t}f^{\frac{\beta+\gamma-1}{\gamma-1}} + \tfrac{\beta}{\gamma}f' + \Big(r + 0.5\tfrac{1}{1-\gamma}\theta^2\Big)f = 0$$

Choosing $\beta = 1-\gamma$ leads to

$$\tfrac{1-\gamma}{\gamma}e^{-\frac{\alpha}{1-\gamma}t} + \tfrac{1-\gamma}{\gamma}f' + \Big(r + 0.5\tfrac{1}{1-\gamma}\theta^2\Big)f = 0,$$

which can be rewritten as

$$f' = Kf + g,$$

with $K := -\frac{\gamma}{1-\gamma}(r + 0.5\frac{1}{1-\gamma}\theta^2)$ and $g(t) := -e^{\frac{\alpha}{\gamma-1}t}$. This is an inhomgeneous ODE for f. Variations of Constants yields

$$f(t) = f^H(t)\Big(f(T) - \int_t^T \frac{g(s)}{f^H(s)}\,ds\Big),$$

where $f^H(t) = e^{-K(T-t)}$ is the solution of the corresponding homogeneous ODE (see Section 5.3). Hence,

$$\begin{aligned}
f(t) &= e^{-K(T-t)}\Big(1 + \underbrace{\int_t^T e^{-\frac{\alpha}{1-\gamma}s+K(T-s)}\,ds}_{\geq 0}\Big) \\
&= e^{-K(T-t)}\Big\{1 - \tfrac{e^{KT}(1-\gamma)}{\alpha+K(1-\gamma)}\Big(e^{-(\frac{\alpha}{1-\gamma}+K)T} - e^{-(\frac{\alpha}{1-\gamma}+K)t}\Big)\Big\}.
\end{aligned}$$

Especially, $f > 0$. Hence, we get the following candidates for the value function and the optimal strategy:

$$\begin{aligned}
G(t,x) &= \frac{1}{\gamma}x^\gamma f^{1-\gamma}(t), \\
\pi^*(t) &= \frac{1}{1-\gamma}\frac{\mu-r}{\sigma^2}, \\
c^*(t) &= \frac{X^*(t)}{f(t)}e^{-\frac{\alpha}{1-\gamma}t}.
\end{aligned} \tag{38}$$

Remark.

- Our candidate G is indeed a $C^{1,2}$-function. Since $\gamma < 1$, we also get $G_{xx}(t,x) = (\gamma-1)x^{\gamma-1}f(t) < 0$.
- Clearly, it needs to be shown that these candidates are indeed the value function and the optimal strategy! This can be done analogously to Proposition 5.4. For the convenience of the reader, the proof of this verification result is also presented.

Proposition 5.6 (Verification Result)
Assume that the coefficients are constant and that the investor has a power utility function. Then the candidates (38) are the value function and the optimal strategy of the portfolio problem (35), where for $\gamma < 0$ only bounded portfolio strategies π are considered.

Proof. Let $(\pi, c) \in \mathcal{A}$. Since $G \in C^{1,2}$, Ito's formula yields

$$\begin{aligned}
&G(T, X^{\pi,c}(T)) + \int_t^T e^{-\alpha s}\tfrac{1}{\gamma}c_s^\gamma\,ds \\
&= G(t,x) + \int_t^T G_t(s, X_s^{\pi,c})ds + \int_t^T G_x(s, X_s^{\pi,c})dX_s^{\pi,c} \\
&\quad +0.5\int_t^T G_{xx}(s, X_s^{\pi,c})d < X^{\pi,c} >_s + \int_t^T \tfrac{1}{\gamma}e^{-\alpha s}c_s^\gamma\,ds \\
&= G(t,x) + \int_t^T \Big(G_t(s, X_s^{\pi,c}) + G_x(s, X_s^{\pi,c})X_s^{\pi,c}(r + \pi_s(\mu - r)) - G_x(s, X_s^{\pi,c})c_s \\
&\qquad +0.5G_{xx}(s, X_s^{\pi,c})(X_s^{\pi,c})^2\pi_s^2\sigma^2 + \tfrac{1}{\gamma}e^{-\alpha s}c_s^\gamma\Big)\,ds \\
&\quad + \int_t^T G_x(s, X_s^{\pi,c})X_s^{\pi,c}\pi_s\sigma\,dW_s
\end{aligned} \tag{39}$$

It is straightforward to show that $\mathrm{E}^{t,x}[\int_t^T G_x(s, X_s^*)X_s^*\pi_s^*\sigma\,dW_s] = 0$ (exercise). Besides, G satisfies the HJB, i.e. for $(\pi, c) = (\pi^*, c^*)$

$$\begin{aligned}
&G_t(s, X_s^*) + G_x(s, X_s^*)X_s^*(r + \pi_s^*(\mu - r)) - G_x(s, X_s^*)c_s^* \\
&+0.5G_{xx}(s, X_s^*)(X_s^*)^2(\pi_s^*)^2\sigma^2 + \tfrac{1}{\gamma}e^{-\alpha s}(c_s^*)^\gamma = 0,
\end{aligned}$$

we obtain

$$\mathrm{E}^{t,x}[G(T,X^*(T))] + \mathrm{E}^{t,x}\Big[\int_t^T e^{-\alpha s}\tfrac{1}{\gamma}(c_s^*)^\gamma\,ds\Big] = G(t,x). \tag{40}$$

implying

$$\mathrm{E}^{t,x}[e^{-\alpha T}\tfrac{1}{\gamma}X^*(T)^\gamma] + \mathrm{E}^{t,x}\Big[\int_t^T e^{-\alpha s}\tfrac{1}{\gamma}(c_s^*)^\gamma\,ds\Big] = G(t,x)$$

because $G(T,x) = e^{-\alpha T}\frac{1}{\gamma}x^\gamma$.

Finally, we need to show that for all $(\pi,c)\in\mathcal{A}$

$$\mathrm{E}^{t,x}[e^{-\alpha T}\tfrac{1}{\gamma}X^{\pi,c}(T)^\gamma] + \mathrm{E}^{t,x}\Big[\int_t^T e^{-\alpha s}\tfrac{1}{\gamma}c_s^\gamma\,ds\Big] \le G(t,x).$$

Since G satisfies the HJB, for all $(\pi,c)\in\mathcal{A}$ we get

$$\begin{aligned}&G_t(s,X_s^{\pi,c}) + G_x(s,X_s^{\pi,c})X_s^{\pi,c}(r+\pi_s(\mu-r)) - G_x(s,X_s^{\pi,c})c_s\\ &+0.5G_{xx}(s,X_s^{\pi,c})(X_s^{\pi,c})^2(\pi_s)^2\sigma^2 + \tfrac{1}{\gamma}e^{-\alpha s}(c_s)^\gamma \le 0.\end{aligned}$$

Therefore, (39) yields

$$G(T,X^{\pi,c}(T)) + \int_t^T e^{-\alpha s}\tfrac{1}{\gamma}c_s^\gamma\,ds \;\le\; G(t,x) + \int_t^T G_x(s,X_s^{\pi,c})X_s^{\pi,c}\pi_s\sigma\,dW_s \tag{41}$$

and now two cases are distinguished.

1st case: $\boldsymbol{\gamma > 0}$. The right-hand side of (41) is a local martingale in T. Since for $\gamma > 0$ the left-hand side is greater than zero, this is even a supermartingale implying

$$\mathrm{E}^{t,x}[G(T,X^{\pi,c}(T))] + \mathrm{E}^{t,x}\Big[\int_t^T e^{-\alpha s}\tfrac{1}{\gamma}c_s^\gamma\,ds\Big] \le G(t,x) \tag{42}$$

and thus

$$\mathrm{E}^{t,x}[e^{-\alpha T}\tfrac{1}{\gamma}X^{\pi,c}(T)^\gamma] + \mathrm{E}^{t,x}\Big[\int_t^T e^{-\alpha s}\tfrac{1}{\gamma}c_s^\gamma\,ds\Big] \le G(t,x).$$

2nd case: $\boldsymbol{\gamma < 0}$. Since for a bounded strategy π we obtain

$$\mathrm{E}[\int_t^T G_x(s,X_s^\pi)X_s^\pi\pi_s\sigma\,dW_s] = 0,$$

the result of the 1st case also holds for $\gamma < 0$ and bounded admissible strategies. □

Remark. If we set $c \equiv 0$, the previous proof is identical to the proof of Proposition 5.4.

5.5 Infinite Horizon

If the optimal portfolio decision over the life-time of an investor is to be determined, then the investment horizon lies very far into the future. For this reason, it reasonable to approximate this situation by a portfolio problem with $T = \infty$.

Hence, the investor's goal function reads

$$\sup_{(\pi,c)\in\mathcal{A}} \mathrm{E}\Big[\int_0^\infty U_1(t,c(t))\,dt\Big], \tag{43}$$

where

$$\mathcal{A} := \Big\{(\pi,c) : (\pi,c) \text{ admissible and } \mathrm{E}\Big[\int_0^\infty U_1(t,c(t))^-\,dt\Big] < \infty\Big\}$$

and

$$\begin{aligned} dX^{\pi,c}(t) &= X^{\pi,c}(t)\Big[(r(t)+\pi(t)'(\mu(t)-r(t)\mathbf{1}))dt + \pi(t)'\sigma dW(t)\Big] - c(t)dt, \\ X^{\pi,c}(0) &= x_0. \end{aligned}$$

As in the previous section we make the following

Assumption 1. The coefficients r, μ, σ are constant and $I = m = 1$.

Additionally, we make the

Assumption 2. $U_1(t,c(t)) = e^{-\alpha t}U(c(t))$, where $\alpha \geq 0$ and U is a utility function.

The value function is given by

$$G(t,x) := \sup_{(\pi,c)\in\mathcal{A}} \mathrm{E}^{t,x}\Big[\int_t^\infty e^{-\alpha s}U(c(s))\,ds\Big].$$

Lemma 5.7
The value function has the representation

$$G(t,x) = e^{-\alpha t}H(x),$$

where H is a function depending on wealth only.

Proof. We obtain

$$\begin{aligned} G(t,x)e^{\alpha t} &= \sup_{(\pi,c)\in\mathcal{A}} \mathrm{E}^{t,x}\Big[\int_t^\infty e^{-\alpha(s-t)}U(c(s))\,ds\Big] \\ &= \sup_{(\pi,c)\in\mathcal{A}} \mathrm{E}^{0,x}\Big[\int_0^\infty e^{-\alpha s}U(c(s))\,ds\Big] =: H(x) \end{aligned}$$

□

Hence, the HJB (37) can be rewritten as follows:

$$\begin{aligned} &\sup_{\pi,c}\Big\{e^{-\alpha t}U(c) + G_t(t,x) + G_x(t,x)x(r+\pi(\mu-r)) - G_x(t,x)c \\ &\qquad +0.5G_{xx}(t,x)x^2\pi^2\sigma^2\Big\} = \\ &\sup_{\pi,c}\Big\{U(c) - \alpha H(x) + H_x(x)x(r+\pi(\mu-r)) - H_x(x)c + 0.5H_{xx}(x)x^2\pi^2\sigma^2\Big\} = 0 \end{aligned}$$

Example "Power Utility". Let $U(c) = \frac{1}{\gamma}c^\gamma$. Then the FOCs read

$$\begin{aligned} H_x(x)x(\mu-r) + H_{xx}(x)x^2\pi(t)\sigma^2 &= 0 \\ c^{\gamma-1} - H_x(x) &= 0 \end{aligned}$$

implying

$$\begin{aligned}\pi^*(t) &= -\frac{\mu - r}{\sigma^2}\frac{H_x(x)}{xH_{xx}(x)}\\ c^*(t) &= H_x^{\frac{1}{\gamma-1}}(x).\end{aligned}$$

Substituting (π^*, c^*) into the HJB leads to

$$\tfrac{1-\gamma}{\gamma}H_x^{\frac{\gamma}{\gamma-1}}(x) - \alpha H(x) + rxH_x(x) - 0.5\theta^2\frac{H_x^2(x)}{H_{xx}(x)} = 0.$$

We try the ansatz $H(x) = \frac{1}{\gamma}x^\gamma k^\beta$ for some constants k and β. Simplifying yields

$$\tfrac{1-\gamma}{\gamma}k^{\frac{\beta}{\gamma-1}} - \tfrac{\alpha}{\gamma} + r + 0.5\tfrac{1}{1-\gamma}\theta^2 = 0$$

Choosing $\beta = \gamma - 1$ leads to

$$k = \tfrac{\gamma}{1-\gamma}\Big(\tfrac{\alpha}{\gamma} - r - 0.5\tfrac{1}{1-\gamma}\theta^2\Big).$$

Our ansatz is only reasonable if $k \geq 0$. For $\gamma < 0$ this is valid in any case. For $\gamma > 0$ we need to assume that

$$\alpha \geq \gamma(r + 0.5\tfrac{1}{1-\gamma}\theta^2). \tag{44}$$

Hence, we get the following candidates for the value function and the optimal strategy:

$$\begin{aligned}G(t,x) &= \frac{1}{\gamma}e^{-\alpha t}x^\gamma k^{\gamma-1},\\ \pi^*(t) &= \frac{1}{1-\gamma}\frac{\mu-r}{\sigma^2},\\ c^*(t) &= k\,X^*(t).\end{aligned} \tag{45}$$

Remark. In contrast to the previous section, the investor now consumes a **time-independent** constant fraction of wealth.

Proposition 5.8 (Verification Result)
Assume that the coefficients are constant and that the investor has a power utility function where for $\gamma < 0$ only bounded portfolio strategies π are considered. If (44) holds as strict inequality, i.e.

$$\alpha > \gamma(r + 0.5\tfrac{1}{1-\gamma}\theta^2), \tag{46}$$

then the candidates (45) are the value function and the optimal strategy of the portfolio problem (43).

Proof. Let $(\pi, c) \in \mathcal{A}$. Analogously to the proof of Proposition 5.6, we get the relations (42) and (40) for the candidate G for the value function of this subsection given by (45)

$$\begin{aligned}\mathrm{E}^{t,x}[G(T, X^{\pi,c}(T))] + \mathrm{E}^{t,x}\Big[\int_t^T e^{-\alpha s}\tfrac{1}{\gamma}c_s^\gamma\,ds\Big] &\leq G(t,x)\\ \mathrm{E}^{t,x}[G(T, X^*(T))] + \mathrm{E}^{t,x}\Big[\int_t^T e^{-\alpha s}\tfrac{1}{\gamma}(c_s^*)^\gamma\,ds\Big] &= G(t,x).\end{aligned}$$

Firstly note that, by the monotonous convergence theorem,[31] for any consumption process c

$$\lim_{T\to\infty} \mathrm{E}^{t,x}\Big[\int_t^T e^{-\alpha s} c_s^\gamma\, ds\Big] = \mathrm{E}^{t,x}\Big[\int_t^\infty e^{-\alpha s} c_s^\gamma\, ds\Big]. \tag{47}$$

Secondly, since $G(T, X^*(T)) = \frac{1}{\gamma} e^{-\alpha T} X^*(T)^\gamma k^{\gamma-1}$, we get

$$\begin{aligned}
\mathrm{E}^{t,x}[G(T, X^*(T))] &= \tfrac{1}{\gamma} k^{\gamma-1}\, e^{-\alpha T} \mathrm{E}^{t,x}[X^*(T)^\gamma] \\
&= \tfrac{1}{\gamma} k^{\gamma-1}\, e^{-\alpha T} x^\gamma \mathrm{E}^{t,x}\Big[e^{\gamma\{r+\frac{1}{1-\gamma}\theta^2 - 0.5\frac{1}{(1-\gamma)^2}\theta^2 - k\}T + \frac{\gamma}{1-\gamma}\theta W(T-t)}\Big] \\
&= \tfrac{1}{\gamma} k^{\gamma-1} x^\gamma e^{-\alpha T + \gamma\{r+\frac{1}{1-\gamma}\theta^2 - 0.5\frac{1}{(1-\gamma)^2}\theta^2 - k + 0.5\frac{\gamma^2}{(1-\gamma)^2}\theta^2\}T} \\
&= \tfrac{1}{\gamma} k^{\gamma-1} x^\gamma e^{-kT} \longrightarrow 0 \quad \text{as } T\to\infty, \qquad (48)
\end{aligned}$$

since (46) ensures that $k > 0$. Therefore,

$$\mathrm{E}^{t,x}\Big[\int_t^\infty e^{-\alpha s} \tfrac{1}{\gamma}(c_s^*)^\gamma\, ds\Big] = G(t,x).$$

To prove

$$\mathrm{E}^{t,x}\Big[\int_t^\infty e^{-\alpha s} \tfrac{1}{\gamma} c_s^\gamma\, ds\Big] \le G(t,x) \tag{49}$$

two cases are distinguished.

1st case: $\gamma > 0$. Since in this case $G(T, X^{\pi,c}(T)) \ge 0$, the relation (49) follows immediately from (42).

2nd case: $\gamma < 0$. Since in this case $G < 0$, the proof is more involved. From (42) it follows that

$$\sup_{\pi,c}\Big\{\mathrm{E}^{t,x}[G(T, X^{\pi,c}(T))]\Big\} + \mathrm{E}^{t,x}\Big[\int_t^T e^{-\alpha s} \tfrac{1}{\gamma} c_s^\gamma\, ds\Big] \le G(t,x)$$

and we obtain the following estimate

$$\begin{aligned}
0 &\ge \sup_{\pi,c}\Big\{\mathrm{E}^{t,x}[G(T, X^{\pi,c}(T))]\Big\} = \sup_{\pi,c}\Big\{\mathrm{E}^{t,x}\Big[\tfrac{1}{\gamma}\, e^{-\alpha T} X^{\pi,c}(T)^\gamma\Big]\Big\}\, k^{\gamma-1} \\
&\overset{(*)}{\ge} \sup_{\pi,c}\Big\{\mathrm{E}^{t,x}\Big[\tfrac{1}{\gamma}\, e^{-\alpha T} X^{\pi,c}(T)^\gamma + \int_t^T e^{-\alpha s} \tfrac{1}{\gamma} c_s^\gamma\, ds\Big]\Big\}\, k^{\gamma-1} \\
&\ge \sup_{\pi}\Big\{\mathrm{E}^{t,x}\Big[\tfrac{1}{\gamma}\, e^{-\alpha T} X^{\pi,0}(T)^\gamma\Big]\Big\}\, k^{\gamma-1} = \sup_{\pi}\Big\{\mathrm{E}^{t,x}\Big[\tfrac{1}{\gamma} X^{\pi,0}(T)^\gamma\Big]\Big\}\, e^{-\alpha T}\, k^{\gamma-1} \\
&\overset{(+)}{=} \tfrac{1}{\gamma} x^\gamma e^{\gamma(r+0.5\frac{1}{1-\gamma}\theta^2)(T-t)}\, e^{-\alpha T} k^{\gamma-1} \overset{(46)}{\longrightarrow} 0 \quad \text{as } T\to\infty,
\end{aligned}$$

where $(*)$ is valid because $\int_t^T e^{-\alpha s}\frac{1}{\gamma} c_s^\gamma\, ds \le 0$ and $(+)$ follows from Proposition 5.4. This establishes (49). □

Remarks.

- The relation

$$\lim_{T\to\infty} \mathrm{E}^{t,x}[G(T, X^{\pi,c}(T))] = 0$$

is the so-called **transversality condition**.

[31] We emphasize that $1/\gamma$ is deliberately omitted because without this factor the argument of the expected value is positive.

- For $\gamma > 0$ the property (48) can be proved by using a different argument:

$$\begin{aligned}
0 &\leq \mathrm{E}^{t,x}[G(T, X^*(T))] = \mathrm{E}^{t,x}[\tfrac{1}{\gamma} X^*(T)^\gamma]\, e^{-\alpha T} k^{\gamma-1} \\
&\leq \mathrm{E}^{t,x}[\tfrac{1}{\gamma} X^{\pi^*,0}(T)^\gamma]\, e^{-\alpha T} k^{\gamma-1} \\
&\leq \sup_\pi \mathrm{E}^{t,x}[\tfrac{1}{\gamma} X^{\pi,0}(T)^\gamma]\, e^{-\alpha T} k^{\gamma-1} \\
&\overset{(+)}{=} \tfrac{1}{\gamma} x^\gamma e^{\gamma(r+0.5\frac{1}{1-\gamma}\theta^2)(T-t)}\, e^{-\alpha T} k^{\gamma-1} \overset{(46)}{\longrightarrow} 0 \quad \text{as } T \to \infty
\end{aligned}$$

  where (+) follows from Proposition 5.4. Applying this argument one can avoid some tedious calculations.
- As in Proposition 5.4, the boundedness assumption for $\gamma < 0$ can be weakened.

References

Bauer, H. (2001): *Measure and integration theory*, De Gruyter, Berlin.

Bingham, N. H.; R. Kiesel (2004): *Risk-neutral valuation*, 2nd ed., Springer, Heidelberg.

Björk, T. (2009): *Arbitrage theory in continuous time*, 3rd ed., Oxford University Press, Oxford.

Cochrance, J. H. (2001): *Asset pricing*, Princeton University Press, Princeton.

Dax, A. (1997): An elementary proof of Farkas' lemma, *SIAM Review* 39, 503-507.

Duffie, D. (2001): *Dynamic asset pricing theory*, 3rd ed., Princeton University Press, Princeton.

Fristedt, B.; L. Gray (1997): *A modern approach to probability theory*, Birkhaeuser Verlag, Basel.

Karatzas, I.; S. E. Shreve (1991): *Brownian motion and stochastic calculus*, 2nd ed., Springer, New York.

Karatzas, I.; S. E. Shreve (1999): *Methods of mathematical finance*, Springer, New York.

Korn, R. (1997): *Optimal portfolios*, World Scientific, Singapore.

Korn, R.; E. Korn (2001): *Option pricing and portfolio optimization - Modern methods of financial mathematics*, AMS, Providence, Rhode Island.

Mangasarian, O. L. (1969): *Nonlinear programming*, New York.

Merton, R. C. (1990): *Continuous-time finance*, Basil Blackwell, Cambridge MA.

Munk, C. (2013): *Financial asset pricing theory*, Oxford University Press, Oxford.

Protter, P. (2005): *Stochastic integration and differential equations*, 2nd ed., Springer, New York.